

# Nonlinear Eigenvalue Problems: A Challenge for Modern Eigenvalue Methods

Volker Mehrmann\*<sup>†</sup>      Heinrich Voss<sup>‡</sup>

November 29, 2004

## Abstract

We discuss the state of the art in numerical solution methods for large scale polynomial or rational eigenvalue problems. We present the currently available solution methods such as the Jacobi-Davidson, Arnoldi or the rational Krylov method and analyze their properties. We briefly introduce a new linearization technique and demonstrate how it can be used to improve structure preservation and with this the accuracy and efficiency of linearization based methods. We present several recent applications where structured and unstructured nonlinear eigenvalue problems arise and some numerical results.

**Keywords.** matrix polynomial, projection method, Krylov-subspace method, Arnoldi method, rational-Krylov method, linearization, structure preservation.

**AMS subject classification.** 65F15, 15A18, 35P30

## 1 Introduction

We discuss numerical methods for the solution of large scale *nonlinear eigenvalue problems*

$$\mathcal{F}(\lambda)x = \mathcal{F}(\lambda; M_0, \dots, M_k, p)x = 0, \quad (1)$$

where for  $\mathbb{F} = \mathbb{C}$  or  $\mathbb{F} = \mathbb{R}$

$$\mathcal{F} : \mathbb{D} \rightarrow \mathbb{F}^{m,n}.$$

---

\*Institut für Mathematik, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG; mehrmann@math.tu-berlin.de

<sup>†</sup>Partially supported by *DFG Research Center MATHEON 'Mathematics for Key Technologies' in Berlin.*

<sup>‡</sup>Arbeitsbereich Mathematik, Technische Universität Hamburg-Harburg, D-21071 Hamburg, FRG; voss@tu-harburg.de

is a family of matrices depending on a variable  $\lambda \in \mathbb{D}$ , where  $\mathbb{D} \subset \mathbb{F}$  is an open set. As in the linear case,  $\lambda \in \mathbb{D}$  is called an eigenvalue of problem (1) if equation (1) has a nontrivial solution  $x \neq 0$ . Then  $x$  is called an eigenvector corresponding to  $\lambda$ .

The function  $\mathcal{F}$  typically depends on some coefficient matrices  $M_0, \dots, M_k \in \mathbb{F}^{m,n}$  and often also on a vector of parameters  $p \in \mathbb{C}^r$ , e.g. material parameters or excitation frequencies. In many applications the purpose of the solution of the eigenvalue problem is to optimize certain properties of the eigenvalues, eigenvectors or the underlying dynamical system with respect to these parameters.

Nonlinear eigenvalue problems arise in a variety of applications. The most widely studied class in applications is the quadratic eigenvalue problem with

$$\mathcal{F}(\lambda) := \lambda^2 M + \lambda C + K \quad (2)$$

that arises in the dynamic analysis of structures, see [48, 27, 74, 90] and the references therein. Here, typically the stiffness matrix  $K$  and the mass matrix  $M$  are real symmetric and positive (semi-)definite, and the damping matrix is general. Another source for such problems are vibrations of spinning structures yielding conservative gyroscopic systems [21, 47, 39], where  $K = K^T$  and  $M = M^T$  are real positive (semi-)definite, and  $C = -C^T$  is real skew-symmetric. In most applications one is interested in the eigenvalues of smallest real part.

A quadratic problem of slightly different structure arises in the study of corner singularities in anisotropic elastic materials, [3, 4, 5, 44, 52, 63, 89], where the problem has the form

$$(\lambda^2 M(p) + \lambda G(p) + K(p))x = 0, \quad (3)$$

with large and sparse coefficient matrices  $M(p) = M(p)^T$ ,  $G(p) = -G(p)^T$ ,  $K(p) = K(p)^T$  that are resulting from a finite element discretization. Here  $M(p)$  and  $-K(p)$  are positive definite and the coefficient matrices depend on a set of material and geometry parameters  $p$  which are varied. The part of the spectrum that is desired are the eigenvalues nearest to the imaginary axis and these are also the eigenvalues for which the error estimates in the finite element discretizations is most favorable.

Polynomial eigenvalue problems of the form (3) are called *even*, see [55], since replacing  $\lambda$  by  $-\lambda$  and transposing gives the same problem. The spectrum of these problems has the *Hamiltonian eigensymmetry*, i.e. it is symmetric with respect to the real and imaginary axis.

Another recently studied structured eigenvalue problem arises in the optimization of the acoustic emissions of high speed trains [34, 35, 55]. The model for the vibrations of the rails leads to a rational eigenvalue problem of the form

$$(\lambda M_1(\omega) + M_0(\omega) + \frac{1}{\lambda} M_1^T(\omega))x = 0, \quad (4)$$

where the coefficients  $M_0, M_1$  are large and sparse complex matrices depending on the excitation frequency  $\omega$ . Here  $M_1(\omega)$  is highly rank deficient and  $M_0(\omega)$  is complex symmetric. The eigenvalues occur in pairs  $\lambda, \frac{1}{\lambda}$  and most of the eigenvalues are at 0 and  $\infty$ . What is needed in the industrial application are the finite nonzero eigenvalues and the corresponding eigenvectors. None of the classical methods worked for this problem and only special methods that were able to deal with the specific structure were able to deliver sufficiently accurate eigenvalue approximations.

Eigenvalue problems of the form (4) (or rather their polynomial representation which is obtained by multiplying (4) by  $\lambda$ ) are called *palindromic* in [55], since transposing and reversing the order yields the same problem. The spectrum has the *symplectic eigensymmetry*, i.e. it is symmetric with respect to the unit circle.

There are many other applications leading to structured or unstructured quadratic eigenvalue problems. A detailed survey has recently been given in [100].

Quadratic eigenvalue problems are special cases of polynomial eigenvalue problems

$$\mathcal{P}(\lambda)x = \left( \sum_{j=0}^k \lambda^j M_j \right) x = 0, \quad (5)$$

with coefficients  $M_j \in \mathbb{F}^{n,n}$ . An important application of polynomial eigenvalue problems is the solution of the optimal control problem to minimize the cost functional

$$\int_{t_0}^{t_1} \sum_{i=0}^k \left[ (q^{(i)})^T Q_i q^{(i)} + u^T R u \right] dt$$

with  $Q_i = Q_i^T$  positive semidefinite,  $R = R^T$  positive definite, subject to the  $k$ -th order control system

$$\sum_{i=0}^k M_i q^{(i)} = B u(t),$$

with control input  $u(t)$  and initial conditions

$$q^{(i)}(t_0) = q_{i,0}, \quad i = 0, 1, \dots, k-1. \quad (6)$$

Application of the linear version of the Pontryagin maximum principle, e.g. [61], leads to the boundary value problem of Euler-Lagrange equations

$$\begin{aligned} \sum_{j=1}^{k-1} \begin{bmatrix} (-1)^{j-1} Q_j & M_{2j}^T \\ M_{2j} & 0 \end{bmatrix} \begin{bmatrix} q^{(2j)} \\ \mu^{(2j)} \end{bmatrix} + \\ \sum_{j=1}^{k-1} \begin{bmatrix} 0 & -M_{2j+1}^T \\ M_{2j+1} & 0 \end{bmatrix} \begin{bmatrix} q^{(2j+1)} \\ \mu^{(2j+1)} \end{bmatrix} + \\ \begin{bmatrix} -Q_0 & M_0^T \\ M_0 & -BR^{-1}B^T \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} = 0, \end{aligned}$$

with initial conditions (6) and  $\mu^{(i)}(t_1) = 0$  for  $i = 0, \dots, k-1$ , where we have introduced the new coefficients  $M_{k+1} = M_{k+2} = \dots = M_{2k} = 0$ . Here, all coefficients of derivatives higher than  $k$  are singular and one obtains a boundary value problem with coefficient matrices that alternate between real symmetric and skew-symmetric matrices. The solution of this boundary value problem can then be obtained by decoupling the forward and backward integration, i.e. by computing the deflating subspace associated with the eigenvalues in the left (or right) half plane, see e.g. [61]. The associated matrix polynomial is even or *odd* depending on the degree and whether the leading coefficient is symmetric or skew-symmetric and the spectrum has the *Hamiltonian eigensymmetry*.

For this problem, even though it is large and sparse, the solution of the boundary value problem requires the computation of a deflating subspace associated with half of the eigenvalues. This can only be done for medium size problems, where it is possible to store the full matrix. If the system size is bigger, then alternative techniques based on low rank approximations to Riccati equations have to be applied, see e.g. [14, 54, 71].

Other polynomial eigenvalue problems of higher degree than two arise when discretizing linear eigenproblems by dynamic elements [74, 103, 104] or by least squares elements [78, 79], i.e. if one uses ansatz functions in a Rayleigh–Ritz approach which depend polynomially on the eigenparameter.

Rational eigenproblems

$$\mathcal{R}(\lambda)x = -Kx + \lambda Mx + \sum_{j=1}^k \frac{\lambda}{\sigma_j - \lambda} C_j x = 0, \quad (7)$$

where  $K = K^T$  and  $M = M^T$  are positive definite and  $C_j = C_j^T$  are matrices of small rank, occur in the study of the free vibration of plates with elastically attached masses [59, 96, 106] or vibrations of fluid solid structures [17, 73, 108].

A similar problem

$$\mathcal{R}(\lambda)x = -Kx + \lambda Mx + \lambda^2 \sum_{j=1}^k \frac{1}{\omega_j - \lambda} C_j x = 0 \quad (8)$$

arises when a generalized linear eigenproblem is condensed exactly [72, 102]. Both these problems have real eigenvalues which can be characterized as min-max values of a Rayleigh functional [106], and in both cases one is interested in a small number of eigenvalues at the lower end of the spectrum or which are close to an excitation frequency.

Another type of rational eigenproblem is obtained for the free vibrations of a structure if one uses a viscoelastic constitutive relation to describe the behavior of a material [31, 32]. A finite element model takes the form

$$\mathcal{R}(\lambda) := \left( \lambda^2 M + K - \sum_{j=1}^k \frac{1}{1 + b_j \lambda} \Delta K_j \right) x = 0, \quad (9)$$

where the stiffness and mass matrices  $K$  and  $M$  are positive definite,  $k$  denotes the number of regions with different relaxation parameters  $b_j$ , and  $\Delta K_j$  is an assemblage of element stiffness matrices over the region with the distinct relaxation constants. Note that the rational problems (4), (7), (8), and (9) can be turned into polynomial eigenvalue problems by multiplying with an appropriate scalar polynomial in  $\lambda$ .

A genuine nonlinear dependence on the eigenparameter appears in dynamic element methods when using non-polynomial ansatz functions [74] or in the stability analysis of vibrating systems under state delay feedback control [37, 38].

Almost all these examples are finite dimensional approximations (typically finite element models) of operator eigenvalue problems and hence are large and sparse. Usually only a small number of eigenvalues in a specific region of the complex plane and associated eigenvectors are of interest.

Furthermore, in most of these applications the eigenvalue problem has extra structure, e.g. is odd, even, or palindromic or a perturbation of a symmetric pencil.

It is common wisdom in numerical analysis, that any kind of extra structure (arising typically from the properties of the underlying physical problem) should be reflected as much as possible in the numerical method. In



Figure 1: Millennium bridge

this way it is guaranteed that the approximate numerical solution properly reflects the physical properties of the system, and also structure preservation typically leads to a gain in efficiency and accuracy. As an example consider problems with Hamiltonian eigensymmetry. It has been shown in [24, 75, 76] that the problem may be well-conditioned under structured perturbations, but ill-posed under unstructured perturbations.

*In summary, many engineering applications lead to large scale polynomial, rational, or more general nonlinear eigenvalue problems with coefficient matrices that are sparse and often have extra structure. Usually only few eigenvalues in a specific region of the complex plane are required. The task of numerical linear algebra is then to design numerical methods that are accurate and efficient for the given problem. The methods should exploit to a maximal extent the sparsity and structure of the coefficient matrices. Furthermore, it should be as accurate as the approximation of the underlying operator problem permits and it should include error and condition estimates.*

Many of the currently used methods in practice do not satisfy these requirements and are not up-to-date. This can e.g. be seen from the Millennium bridge over the river Thames in London, which had to be closed right on opening day, since the step frequency of 0.9 Hz was close to a resonance frequency of the bridge. The cost to build in new dampers was on the order of 5 million pounds. The Millennium bridge example demonstrates that millions are invested in new industrial and architectural designs but

the numerical methods that are used are often very old and not adequate for the current applications.

*More research into new numerical methods and appropriate software for nonlinear eigenvalue problems is urgently needed.*

In linear eigenvalue problems  $Ax = \lambda x$  or  $Ax = \lambda Ex$ , well established methods are available, that include error and condition estimates. These methods are able to deal with most of the small or large scale problems in practice [1, 6, 16, 53, 58] and specific methods have been developed to handle extra structures [9, 12, 61, 63, 64].

*For nonlinear eigenvalue problems, there are essentially no analogous packages that reach the standard of those for linear problems.*

There are several reasons for this lack of adequate software. First of all, the essential tools in most numerically stable eigenvalue methods, the Schur form and generalized Schur form [28], are in general, not available for nonlinear eigenvalue problems. Not even an analogue to the Jordan or Kronecker/Weierstrass form is available, [27]. These missing tools make numerical techniques such as deflation, purging and locking of eigenvalues, or implicit restarts very difficult. Also the sensitivity and round-off error analysis for nonlinear eigenvalue problems is still in its infancy, see e.g. [99, 91].

The theoretical analysis and the numerical methods for polynomial eigenvalue problems usually proceed via *linearization*, i.e. via the embedding of the nonlinear eigenvalue problem into a larger linear eigenvalue problem [27]. This is not ideal, since it makes the problem much bigger and also may significantly increase the conditioning of the problem, i.e. the resulting linear eigenvalue problem may be much more sensitive to perturbations than the original problem, see [99]. Also in the classical linearization approaches usually the inherent symmetry structure of the problem is destroyed. Some progress in the construction of better linearizations comes from the recent work [56], where a systematic new linearization approach has been developed that allows and analyzes also structure preserving linearizations, [33, 55, 64]. We will briefly review this new linearization concept in section 2.

Most of the literature on polynomial eigenvalue problems discusses only *regular matrix polynomials*, i.e. square matrix polynomials, where  $\det(\mathcal{P}(\lambda))$  does not vanish identically. Singular problems, however, arise frequently in practice, when automatic modeling leads to overdetermined systems with redundant equations, [22, 65, 98]. Often, the case of singular matrix polynomials can be reduced to the case of regular matrix polynomials using appropriate reduction procedures that have recently been derived in [62, 91]. But these procedures are not yet applicable to large scale problems and

rather expensive even for small scale problems.

Finally, it is often assumed that the leading coefficient matrix  $M_k$  is nonsingular or even the identity matrix. In many applications, such as constraint multi-body systems [22, 87], circuit simulation [26], optical waveguide design [88], or problem (4), however, the leading coefficient is singular, i.e. the matrix polynomial has eigenvalues at infinity. In such a case not all linearizations properly reflect the multiplicities of the eigenvalue infinity, see [62, 91]. In [49] it has recently been suggested to use *strong linearizations* which properly reflect the multiplicities in linearizations. The new general linearization approach of [56] yields strong linearizations and the reduction procedure of [62, 91] allows to deflate this part of the matrix polynomial. In problems with extra structure, this can be done in a structured way, see e.g. [34, 35, 62].

*More research is needed on the development of methods and the appropriate perturbation and error analysis, using the original data of the problem and not the matrix pencil arising from the linearization.*

Despite the urgent need for more research and many of the described difficulties, there exist many classical but also several new and promising numerical methods for nonlinear eigenvalue problems. We will review the state of the art in Section 3 and present some numerical examples in Section 4.

## 2 Linearization

The classical approach to solve  $k$ -th degree polynomial eigenvalue problems of the form (5) or rational eigenvalue problems of the forms (7)–(9) is to first perform a *linearization* [27], i.e. to transform the problem to an equivalent linear eigenvalue problem  $(\lambda E - A)x = 0$  with the same eigenvalues.

The transformation between the polynomials and the linear pencil is performed with the help of *unimodular* matrix polynomials, i.e., matrix polynomials  $Q(\lambda)$  such that  $\det Q(\lambda)$  is a nonzero constant, independent of  $\lambda$ .

**Definition 2.1 (Linearization [27])** *Let  $\mathcal{P}(\lambda)$  be a matrix polynomial of  $k$ -th degree and coefficients in  $\mathbb{F}^{n,n}$ . A pencil  $\mathcal{L}(\lambda) = \lambda E - A$  with  $E, A \in \mathbb{F}^{kn, kn}$  is called a linearization of  $\mathcal{P}(\lambda)$  if there exist unimodular matrix polynomials  $Q_1(\lambda), Q_2(\lambda)$  such that*

$$Q_1(\lambda)\mathcal{L}(\lambda)Q_2(\lambda) = \left[ \begin{array}{c|c} \mathcal{P}(\lambda) & 0 \\ \hline 0 & I_{(k-1)n} \end{array} \right].$$



The most commonly used linearization are the first companion form

$$\left( \lambda \begin{bmatrix} M_k & 0 & \cdots & 0 \\ 0 & I_n & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I_n \end{bmatrix} + \begin{bmatrix} M_{k-1} & M_{k-2} & \cdots & M_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{bmatrix} \right) \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_1 \end{bmatrix} = 0,$$

which is obtained by introducing the new vectors  $x_i = \lambda^{i-1}x$ ,  $i = 1, \dots, k$ , and the analogous second companion form with the pencil

$$\left( \lambda \begin{bmatrix} M_k & 0 & \cdots & 0 \\ 0 & I_n & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I_n \end{bmatrix} + \begin{bmatrix} M_{k-1} & -I_n & \cdots & 0 \\ M_{k-2} & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & -I_n \\ M_0 & 0 & \cdots & 0 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = 0.$$

But the companion forms are not the only possible linearizations. In a recent paper [56] a systematic approach has been presented to generate large classes of linearizations by constructing  $kn \times kn$  matrix pencils  $\mathcal{L}(l) = lX + Y$  with the properties

$$\mathcal{L}(l) \begin{bmatrix} l^{k-1}I_n \\ \vdots \\ lI_n \\ I_n \end{bmatrix} = \begin{bmatrix} v_1P(l) \\ \vdots \\ v_{k-1}P(l) \\ v_kP(l) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} l^{k-1}I_n \\ \vdots \\ lI_n \\ I_n \end{bmatrix}^T \mathcal{L}(\lambda) = \begin{bmatrix} v_1\mathcal{P}(l) \\ \vdots \\ v_{k-1}\mathcal{P}(l) \\ v_k\mathcal{P}(l) \end{bmatrix}^T \quad (10)$$

for some  $v = [v_1, \dots, v_k]^T \in \mathbb{F}^k$ . It has been shown in [56] that in order to properly reflect the infinite eigenvalues (which occur if the leading coefficient  $M_k$  is singular) one should consider matrix pencils  $\mathcal{L}(\lambda)$  that satisfy one or both conditions in (10). For pencils that satisfy both conditions, the following Theorem classifies when such pencils are linearizations.

**Theorem 2.2** [56] *Suppose that  $\mathcal{P}(l)$  is a regular matrix polynomial, i.e.  $\mathcal{P}(l)$  is square and its determinant does not vanish identically, and suppose that  $\mathcal{L}(l)$  satisfies both conditions in (10). Define the scalar polynomial  $q(\xi) = \sum_{i=1}^k v_i \xi^{k-i} + v_k$ . Then  $\mathcal{L}(l)$  is a linearization for  $\mathcal{P}(l)$  if and only if the sets  $\{\text{Roots of } q(\xi)\}$  and  $\{\text{Eigenvalues of } \mathcal{P}(l)\}$  are disjoint. Here it is to be understood that  $\infty$  may also be an element of either set.*

Using this theorem, many different linearizations can be constructed, in particular linearizations that have the same structure as the original matrix polynomial. Furthermore, the freedom in the choice of the coefficients may be used to make the resulting linear eigenvalue problem as well conditioned as possible, although at the writing of this paper it is an open problem how to do this.

**Example 2.3** [3, 63] Considering an even quadratic eigenvalue problem of the form (3) and the even polynomial  $q(\xi) = 1$ , we obtain the linearization

$$\mathcal{L}(\lambda) = \lambda \begin{bmatrix} 0 & -M \\ M & G \end{bmatrix} + \begin{bmatrix} M & 0 \\ 0 & K \end{bmatrix}.$$

Since  $q(\xi)$  has only the eigenvalue  $\infty$ , we have by Theorem 2.2 that  $\mathcal{L}(\lambda)$  is a linearization if and only if the mass matrix  $M$  is nonsingular.

**Example 2.4** [34, 35, 55] Consider a palindromic quadratic eigenvalue problem of the form (4) and the palindromic polynomial  $q(\xi) = \xi + 1$ , we obtain the palindromic pencil  $\lambda Z + Z^T$  with

$$Z = \begin{bmatrix} M_1 & M_0 - M_1^T \\ M_1 & M_1 \end{bmatrix}$$

which is by Theorem 2.2 a linearization if and only if  $-1$  is not an eigenvalue of (4). (This condition is always satisfied in the industrial application, since in this example  $-1$  corresponds to an undamped vibration of the rail.)

**Example 2.5** Consider a symmetric eigenvalue problem of the form (2), i.e.  $M, C, K$  are symmetric, and the general polynomial  $q(\xi) = v_1\xi + v_2$ . A simple calculation yields pencils of the form

$$\lambda \begin{bmatrix} v_1M & v_2M \\ v_2M & v_2C - v_1K \end{bmatrix} + \begin{bmatrix} v_1C - v_2M & v_1K \\ v_1K & v_2K \end{bmatrix}.$$

By Theorem 2.2 this is a linearization if and only if no eigenvalue of (2) is a root of  $q(\xi)$ .

### 3 Numerical methods

There is a vast literature on numerical methods for nonlinear eigenvalue problems. In general, one has to distinguish between dense and large sparse problems.

For dense problems, the size of the problems that can be treated with numerical methods is limited to a few thousand depending on the available storage capacities. Methods for small dense problems, however, are needed in most of the iterative projection methods for large sparse problems. These iterative projection methods make use of the sparse matrix structure and typically require only matrix vector multiplication with the coefficient matrices plus possibly sparse factorizations of matrices, when shift-and-invert is used to get eigenvalues in the interior of the spectrum. Again here the available storage sets the limit for the system sizes that can be dealt with. Using the sparsity and the symmetry structure of the coefficient matrices, nowadays problems of size on the order of  $n = 10^7$  can be treated.

In the polynomial or rational case the simplest approach is to use linearization and then to apply standard methods for linear eigenvalue problems, as they are available in [1, 13, 16, 53, 58]. Despite the fact that this increases the dimension, it is usually the method of choice if no special properties of the problem are known. In the case of structured problems the linearization should reflect the structure and then a structure preserving method for the linearized problem should be used. It is currently under discussion to improve the features of [1, 16, 58] to make better use of the resulting structures in the linearized problems.

### 3.1 Newton type methods and inverse iteration

For general nonlinear eigenvalue problems, the classical approach is to formulate the eigenvalue problem as a system of nonlinear equations and to use variations of Newton's method or the inverse iteration method. For the characteristic equation

$$\det \mathcal{F}(\lambda) = 0, \quad (11)$$

it was suggested in [45, 46] to use a  $QR$ -decomposition with column pivoting  $\mathcal{F}(\lambda)P(\lambda) = Q(\lambda)R(\lambda)$ , where  $P(\lambda)$  is a permutation matrix which is chosen such that the diagonal elements  $r_{jj}(\lambda)$  of  $R(\lambda)$  are decreasing in magnitude, i.e.  $|r_{11}(\lambda)| \geq |r_{22}(\lambda)| \geq \dots \geq |r_{nn}(\lambda)|$ . Then  $\lambda$  is an eigenvalue if and only if  $r_{nn}(\lambda) = 0$ .

Applying Newton's method to this equation, one obtains the iteration

$$\lambda_{k+1} = \lambda_k - \frac{1}{e_n^H Q(\lambda_k)^H \mathcal{F}'(\lambda_k) P(\lambda_k) R(\lambda_k)^{-1} e_n} \quad (12)$$

for approximations to an eigenvalue of problem (11). (Here  $e_n$  denotes the  $n$ -th unit vector.) Approximations to left and right eigenvectors can be

obtained from

$$y_k = Q(\lambda_k)e_n \quad \text{and} \quad x_k = P(\lambda_k)R(\lambda_k)^{-1}e_n.$$

An improved version of this method was suggested in [40, 41] and also quadratic convergence was shown. A similar approach was presented in [116], via a representation of Newton's method using the  $LU$ -factorization of  $\mathcal{F}(\lambda)$ . Other variations of this method can be found in [117, 118]. However, this relatively simple idea is not efficient, since it computes eigenvalues one at a time and needs several  $\mathcal{O}(n^3)$  factorizations per eigenvalue. It is, however, useful in the context of iterative refinement of computed eigenvalues and eigenvectors.

Another method that also solves the purpose of iterative refinement is the *nonlinear version of inverse iteration*. For linear eigenproblems  $Ax = \lambda x$  it is well known that inverse iteration is equivalent to Newton's method applied to the nonlinear system

$$\begin{bmatrix} Ax - \lambda x \\ v^H x - 1 \end{bmatrix} = 0$$

where  $v \in \mathbb{C}^n$  is suitably chosen. Correspondingly, for the nonlinear problem

$$\begin{bmatrix} \mathcal{F}(\lambda)x \\ v^H x - 1 \end{bmatrix} = 0$$

one step of Newton's method gives

$$\begin{bmatrix} \mathcal{F}(\lambda_k) & \mathcal{F}'(\lambda_k)x_k \\ v^H & 0 \end{bmatrix} \begin{bmatrix} x_{k+1} - x_k \\ \lambda_{k+1} - \lambda_k \end{bmatrix} = - \begin{bmatrix} \mathcal{F}(\lambda_k)x_k \\ v^H x_k - 1 \end{bmatrix}. \quad (13)$$

The first component yields

$$x_{k+1} = -(\lambda_{k+1} - \lambda_k)\mathcal{F}(\lambda_k)^{-1}\mathcal{F}'(\lambda_k)x_k, \quad (14)$$

i.e. the direction of the new approximation to an eigenvector is  $u_{k+1} := \mathcal{F}(\lambda_k)^{-1}\mathcal{F}'(\lambda_k)x_k$ . Assuming that  $x_k$  is already normalized by  $v^H x_k = 1$ , the second component of (13) reads  $v^H x_{k+1} = v^H x_k$ , and multiplying equation (14) by  $v^H$  yields  $\lambda_{k+1} = \lambda_k - \frac{v^H x_k}{v^H u_{k+1}}$ . Hence, for nonlinear eigenproblems inverse iteration takes the form given in Algorithm 1.

This algorithm (being a variant of Newton's method) converges locally and quadratically to  $(x, \lambda)$  [2, 68].

The normalization condition can be modified in each step of inverse iteration. It was suggested in [80] to use  $v_k = \mathcal{F}(\lambda_k)^H y_k$  for the normalization,

---

**Algorithm 1** Inverse iteration

---

- 1: Start with  $\lambda_0, x_0$  such that  $v^H x_0 = 1$
  - 2: **for**  $k = 0, 1, 2, \dots$  until convergence **do**
  - 3:   solve  $\mathcal{F}(\lambda_k)u_{k+1} = \mathcal{F}'(\lambda_k)x_k$  for  $u_{k+1}$
  - 4:    $\lambda_{k+1} = \lambda_k - (v^H x_k)/(v^H u_{k+1})$
  - 5:   normalize  $x_{k+1} = u_{k+1}/v^H u_{k+1}$
  - 6: **end for**
- 

where  $y_k$  is an approximation to a left eigenvector. Then the update for  $\lambda$  becomes

$$\lambda_{k+1} = \lambda_k - \frac{y_k^H \mathcal{F}(\lambda_k)x_k}{y_k^H \mathcal{F}'(\lambda_k)x_k},$$

which is the Rayleigh functional for general nonlinear eigenproblems proposed in [48], and which can be interpreted as one Newton step for solving the equation  $f_k(\lambda) := y_k^H \mathcal{F}(\lambda)x_k = 0$ . For linear Hermitian eigenproblems this gives cubic convergence if  $\lambda_k$  is updated by the Rayleigh quotient [19, 70]. The same is true [78] for symmetric nonlinear eigenproblems having a Rayleigh functional if we replace statement 4 in Algorithm 1 by  $\lambda_{k+1} = p(u_{k+1})$ , where  $p(u_{k+1})$  denotes the real root of  $u_{k+1}^H \mathcal{F}(\lambda)u_{k+1} = 0$  closest to  $\lambda_k$ . In [67] Newton's method is considered for the complex function  $\beta(\lambda)$  defined by

$$\mathcal{F}(\lambda)u = \beta(\lambda)x, \quad s^H u = \kappa,$$

where  $\kappa$  is a given constant, and  $x$  and  $u$  are given vectors. This approach generalizes the method (12), inverse iteration, and a method proposed in [69]. It was proved that the rate of convergence is quadratic, and that cubic convergence can be obtained if not only  $\lambda$ , but also  $x$  and/or  $s$  are updated appropriately, thus unifying the results in [2, 45, 46, 48, 68, 69].

The disadvantage of inverse iteration with respect to efficiency is the large number of factorizations that are needed for each of the eigenvalues. The obvious idea then is to use a version of a simplified Newton method, where the shift  $\sigma$  is kept fixed during the iteration, i.e. to use,  $x_{k+1} = (A - \sigma I)^{-1}x_k$ . However, in general this method does not converge in the nonlinear case. The iteration converges to an eigenpair of a linear problem  $\mathcal{F}(\sigma)x = \gamma \mathcal{F}'(\tilde{\lambda})x$ , from which one cannot recover an eigenpair of the nonlinear problem (1). A remedy against this problem was proposed in [66]. Assuming that  $\mathcal{F}(\lambda)$  is twice continuously differentiable, then Algorithm 1 gives

$$x_k - x_{k+1} = x_k + (\lambda_{k+1} - \lambda_k)\mathcal{F}(\lambda_k)^{-1}\mathcal{F}'(\lambda_k)x_k$$

$$\begin{aligned}
&= \mathcal{F}(\lambda_k)^{-1}(\mathcal{F}(\lambda_k) + (\lambda_{k+1} - \lambda_k)\mathcal{F}'(\lambda_k))x_k \\
&= \mathcal{F}(\lambda_k)^{-1}\mathcal{F}(\lambda_{k+1})x_k + \mathcal{O}(|\lambda_{k+1} - \lambda_k|^2).
\end{aligned}$$

Neglecting the second order term one gets

$$x_{k+1} = x_k - \mathcal{F}(\lambda_k)^{-1}\mathcal{F}(\lambda_{k+1})x_k.$$

The advantage of this approach is that replacing  $\lambda_k$  by a fixed shift  $\sigma$  does not lead to misconvergence. The method can be implemented as in Algorithm 2, see [66].

---

**Algorithm 2** Residual inverse iteration

---

- 1: Let  $v$  be a normalization vector and start with an approximations  $\sigma$  and  $x_1$  to an eigenvalue and corresponding eigenvector of (1) such that  $v^H x_1 = 1$
  - 2: **for**  $k = 1, 2, \dots$  until convergence **do**
  - 3: solve  $v^H \mathcal{F}(\sigma)^{-1} \mathcal{F}(\lambda_{k+1}) x_k = 0$  for  $\lambda_{k+1}$   
or  $x_k^H \mathcal{F}(\lambda_{k+1}) x_k = 0$  if  $\mathcal{F}(\lambda)$  is Hermitian and  $\lambda_{k+1}$  is real
  - 4: compute the residual  $r_k = \mathcal{F}(\lambda_{k+1}) x_k$
  - 5: solve  $\mathcal{F}(\sigma) d_k = r_k$  for  $d_k$
  - 6: set  $z_{k+1} = x_k - d_k$
  - 7: normalize  $x_{k+1} = z_{k+1} / v^H z_{k+1}$
  - 8: **end for**
- 

If  $\mathcal{F}(\lambda)$  is twice continuously differentiable,  $\hat{\lambda}$  a simple zero of  $\det \mathcal{F}(\lambda) = 0$ , and if  $\hat{x}$  is an eigenvector normalized by  $v^H \hat{x} = 1$ , then the residual inverse iteration converges for all  $\sigma$  sufficiently close to  $\hat{\lambda}$ , and one has the estimate

$$\frac{\|x_{k+1} - \hat{x}\|}{\|x_k - \hat{x}\|} = \mathcal{O}(|\sigma - \hat{\lambda}|) \quad \text{and} \quad |\lambda_{k+1} - \hat{\lambda}| = \mathcal{O}(\|x_k - \hat{x}\|^q),$$

where  $q = 2$  if  $\mathcal{F}(\lambda)$  is Hermitian,  $\hat{\lambda}$  is real, and  $\lambda_{k+1}$  solves  $x_k^H \mathcal{F}(\lambda_{k+1}) x_k = 0$  in Step 3, and  $q = 1$  otherwise, see [66].

A variant of this approach is the method of successive linear approximations of [80]. If  $\mathcal{F}$  is twice continuously differentiable, and  $\hat{\lambda}$  is an eigenvalue of problem (1) such that  $\mathcal{F}'(\hat{\lambda})$  is nonsingular and 0 is an algebraically simple eigenvalue of  $\mathcal{F}'(\hat{\lambda})^{-1}\mathcal{F}(\hat{\lambda})$ , then the method in Algorithm 3 converges quadratically to  $\hat{\lambda}$ , see [80].

The discussed versions of inverse iteration apply to general nonlinear eigenproblems, although for Hermitian problems and real eigenvalues they converge faster if the eigenvalue approximations are updated using the Rayleigh

---

**Algorithm 3** Method of successive linear problems

---

- 1: Start with an approximation  $\lambda_1$  to an eigenvalue of (1)
  - 2: **for**  $k = 1, 2, \dots$  until convergence **do**
  - 3:   solve the linear eigenproblem  $\mathcal{F}(\lambda_k)u = \theta\mathcal{F}'(\lambda_k)u$
  - 4:   choose an eigenvalue  $\theta$  smallest in modulus
  - 5:    $\lambda_{k+1} = \lambda_k - \theta$
  - 6: **end for**
- 

functional. For Hermitian problems that allow a variational characterization of their eigenvalues [20, 29, 30, 77, 107, 112], an alternative is to use the *safeguarded iteration*. The method was introduced in [115] for over-damped problems, and in [113] for the non-over-damped case.

Let  $J \subset \mathbb{R}$  be an open interval which may be unbounded, and assume that  $\mathcal{F}(\lambda) \in \mathbb{C}^{n,n}$  is a family of Hermitian matrices, where the elements are differentiable in  $\lambda$ . If one assumes that for every  $x \in \mathbb{C}^n \setminus \{0\}$  the real equation

$$f(\lambda, x) := x^H \mathcal{F}(\lambda)x = 0 \quad (15)$$

has at most one solution  $\lambda \in J$ , then (15) defines a functional  $\rho$  on some subset  $D \subset \mathbb{C}^n$  which generalizes the Rayleigh quotient for linear pencils  $\mathcal{F}(\lambda) = \lambda E - A$ , and which is called *Rayleigh functional of the nonlinear eigenvalue problem* (1). If one assumes further that  $x^H \mathcal{F}'(\rho(x))x > 0$  for every  $x \in D$  (generalizing the definiteness requirement for linear pencils), then by the implicit function theorem  $D$  is an open set, and differentiating the identity  $x^H \mathcal{F}(\rho(x))x = 0$  one obtains that the eigenvectors of (1) are stationary points of  $\rho$ .

Under these conditions in [112] a minimax principle for the nonlinear eigenproblem (1) was proved if the eigenvalues are enumerated appropriately. A value  $\lambda \in J$  is an eigenvalue of (1) if and only if  $\mu = 0$  is an eigenvalue of the matrix  $\mathcal{F}(\lambda)$ , and by Poincaré's max-min principle there exists  $m \in \mathbb{N}$  such that

$$0 = \max_{\dim V=m} \min_{x \in V, x \neq 0} \frac{x^H \mathcal{F}(\lambda)x}{\|x\|^2}.$$

One assigns this  $m$  to  $\lambda$  as *its number* and calls  $\lambda$  an  *$m$ -th eigenvalue of problem (1)*.

Under the above assumptions, it was shown in [112]) that for every  $m \in \{1, \dots, n\}$  problem (1) has at most one  $m$ -th eigenvalue in  $J$ , which can be characterized by

$$\lambda_m = \min_{\dim V=m, D \cap V \neq \emptyset} \sup_{v \in D \cap V} \rho(v). \quad (16)$$

Conversely, if

$$\lambda_m := \inf_{\dim V=m, D \cap V \neq \emptyset} \sup_{v \in D \cap V} \rho(v) \in J,$$

then  $\lambda_m$  is an  $m$ -th eigenvalue of (1), and the characterization (16) holds. The minimum is attained by the invariant subspace of  $\mathcal{F}(\lambda_m)$  corresponding to its  $m$  largest eigenvalues, and the supremum is attained by any eigenvector of  $\mathcal{F}(\lambda_m)$  corresponding to  $\mu = 0$ .

The enumeration of eigenvalues and the fact that the eigenvectors of (1) are the stationary vectors of the Rayleigh functional suggests the following Algorithm 4. It was shown in [111, 113] that the safeguarded iteration has

---

**Algorithm 4** Safeguarded iteration

---

- 1: Start with an approximation  $\sigma_1$  to the  $m$ -th eigenvalue of (1)
  - 2: **for**  $k = 1, 2, \dots$  until convergence **do**
  - 3:   determine an eigenvector  $x_k$  corresponding to the  $m$ -largest eigenvalue of  $\mathcal{F}(\sigma_k)$
  - 4:   solve  $x_k^H \mathcal{F}(\sigma_{k+1}) x_k = 0$  for  $\sigma_{k+1}$
  - 5: **end for**
- 

the following convergence properties.

- (i) If  $\lambda_1 := \inf_{x \in D} \rho(x) \in J$  and  $x_1 \in D$  then the iteration converges globally to  $\lambda_1$ .
- (ii) If  $\lambda_m \in J$  is a  $m$ -th eigenvalue of (1) which is simple, then the iteration converges locally and quadratically to  $\lambda_m$ .
- (iii) Let  $\mathcal{F}(\lambda)$  be twice continuously differentiable, and assume that  $\mathcal{F}'(\lambda)$  is positive definite for  $\lambda \in J$ . If, in Step 3 of Algorithm 4,  $x_k$  is chosen to be an eigenvector corresponding to the  $m$ -th largest eigenvalue of the generalized eigenproblem  $\mathcal{F}(\sigma_k)x = \mu \mathcal{F}'(\sigma_k)x$ , then the convergence is even cubic.

The methods that have been discussed in this subsection are appropriate for problems for which matrix factorizations of  $\mathcal{F}(\sigma)$  can be carried out and the storage for the factors is available. It is an obvious idea to use iterative methods for the solution of the occurring linear systems, but the convergence properties of the methods have to be carefully analyzed in this case. This topic is also discussed in the following section, where methods that can be applied to large sparse problems are discussed.



### 3.2 Iterative projection methods for linear problems

For sparse linear eigenvalue problems  $Ax = \lambda x$ , iterative projection methods like the Lanczos, Arnoldi, rational Krylov or Jacobi–Davidson method are well established. The basic idea of all these methods is the construction of a *search space* (typically a Krylov subspace) followed by projection of the problem into this subspace. This leads to a small dense problem that is then handled by one of the techniques from the previous section and the eigenvalues of the projected problem are used as approximations to the eigenvalues of the large sparse problem. The main feature of all these methods is that matrix factorizations are avoided as much as possible (except in the context of preconditioning) and the generation of the search space is usually done via an iterative procedure that is based on matrix vector products that can be cheaply obtained.

Two basic types of iterative projection methods are in use: The first type consists of methods which expand the subspaces independently of the eigenpair of the projected problem and which use Krylov subspaces of  $A$  or  $(A - \sigma I)^{-1}$  for some shift  $\sigma$ . These methods include the Arnoldi, Lanczos or rational Krylov method. The other type of methods aim at a particular eigenpair and choose the expansion  $q$  such that it has a high approximation potential for a desired eigenvalue/eigenvector or invariant subspace. An example for this approach is the Jacobi–Davidson method.

For the Arnoldi method and similarly for other Krylov subspace methods, the search space is a *Krylov space*

$$\mathcal{K}_k(A, v_1) = \text{span}\{v_1, Av_1, A^2v_1, \dots, A^{k-1}v_1\},$$

where  $v_1$  is an appropriately chosen initial vector. The Arnoldi method produces an orthogonal basis  $V_k$  of  $\mathcal{K}_k(A, v_1)$  such that the projected matrix  $H_k$  (associated with  $\mathcal{K}_k$ ) is upper Hessenberg and satisfies  $AV_k = V_k H_k + f_k e_k^T$ , where  $e_k \in \mathbb{R}^k$  is the  $k$ -th unit vector and  $f_k$  is orthogonal to the columns of  $V_k$ , i.e.  $V_k^H f_k = 0$ . The orthogonality of  $V_k$  implies that  $V_k^H A V_k = H_k$  is the orthogonal projection of  $A$  to  $\mathcal{K}_k(A, v_1)$ .

If  $(y, \theta)$  is an eigenpair of the projected problem, and  $x = V_k y$  is the corresponding approximation to an eigenvector of  $Ax = \lambda x$  (which is called a *Ritz vector* corresponding to the *Ritz value*  $\theta$ ), then the residual satisfies

$$r := Ax - \theta x = AV_k y - \theta V_k y = V_k H_k y - \theta V_k y + f_k e_k^H y = (e_k^H y) f_k.$$

Hence, one obtains an error indicator  $\|r\| = |e_k^T y| \cdot \|f_k\|$  for the eigenpair approximation  $(x, \theta)$  without actually computing the Ritz vector  $x$ . If  $A$  is Hermitian then this is even an error bound.

The Arnoldi method together with its variants, is today a standard solver for sparse linear eigenproblems. It is implemented in the package ARPACK [53] and the MATLAB command `eigs`, see also [6]. The method typically converges to the extreme eigenvalues first. If one is interested in eigenvalues in the interior of the spectrum, or in eigenvalues close to a given focal point  $\tau$ , then one can apply the method in a shift-and-invert fashion, i.e. to the matrix  $(A - \tau I)^{-1}$  or an approximation of it. In this case one has to determine a factorization of  $A - \tau I$ , which, however, may be prohibitive for very large problems.

A way out of this dilemma is the *Jacobi–Davidson method* of [94]. Let  $(x, \theta)$  be an approximation to an eigenpair obtained by a projection method with subspace  $V$ . We assume that  $\|x\| = 1$ ,  $\theta = x^H A x$  and  $r := Ax - \theta x \perp x$ . Then the most desirable orthogonal correction  $z$  solves the equation

$$A(x + z) = \lambda(x + z), \quad z \perp x. \quad (17)$$

As  $z \perp x$ , the operator  $A$  can be restricted to the subspace orthogonal to  $x$  yielding  $\tilde{A} := (I - xx^H)A(I - xx^H)$ , and from  $\theta = x^H A x$  it follows that

$$A = \tilde{A} + Axx^H + xx^H A - \theta xx^H.$$

Hence, from (17) and  $\tilde{A}z = 0$  we obtain that  $(\tilde{A} - \lambda I)z = -r + (\lambda - \theta - x^H A z)x$ . Since both the left hand side and  $r$  are orthogonal to  $x$ , it follows that the factor  $\lambda - \theta - x^H A z$  must vanish, and therefore the correction  $z$  has to satisfy  $(\tilde{A} - \lambda I)z = -r$ . Since  $\lambda$  is unknown, it is replaced by a Ritz approximation  $\theta$ , and one ends up with the correction equation

$$(I - xx^H)(A - \theta I)(I - xx^H)z = -r.$$

It can be shown that the expanded space  $[V, z]$  for the Jacobi–Davidson method contains the direction  $u = (A - \theta I)^{-1}x$ , which is obtained by one step of inverse iteration [93]. One therefore can expect similar approximation properties, i.e. quadratic or even cubic convergence, if the problem is Hermitian.

Obviously, the Jacobi–Davidson method is aiming at a particular eigenvalue (close to the shift  $\theta$ ). If one is interested in more than one eigenvalue, one typically uses a deflation based on a partial Schur decomposition of the matrix  $A$ , see [23].

Both, the shift-and-invert Arnoldi method and the Jacobi–Davidson method have to solve a large linear system. However, while in the Arnoldi method this system in general needs to be solved very accurately to get fast convergence, numerical experiments demonstrate that in the Jacobi–Davidson

method it suffices to solve this system approximately to maintain fast convergence. Typically only a small number of steps of a preconditioned Krylov subspace method are sufficient to obtain a good expansion  $z$  for the search space  $V$ . Implementation details of the Jacobi–Davidson method for various types of linear eigenvalue problems can be found in [6]. Implementations in FORTRAN and MATLAB can be downloaded from <http://www.math.ruu.nl/people/sleijpen>.

Many, but not all of the ideas in these projection methods can be generalized also to nonlinear eigenproblems.

### 3.3 Structure preserving iterative projection methods for linearized problems

For polynomial and rational problems, the easiest approach is to use linearization and to apply the projection method to the linearization. As we have seen in Section 2 one should use structure preserving linearizations, make use of the symmetries in the generation of the search space and also make sure that the projection leads to a small problem of the same structure. Recently several structure-preserving Krylov subspace methods have been developed [3, 10, 11, 63, 64, 114]. Each of these requires that a structured generalized eigenvalue problem, such as the linearizations obtained in Examples 2.3 or 2.4, is reduced further to a Hamiltonian or symplectic matrix. (A  $2n \times 2n$  matrix  $H$  is called *Hamiltonian* if  $(HJ)^T = HJ$  and *symplectic* if  $HJH^T = J$ , where  $J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ .) To see how this can be done very efficiently, consider the even pencil of Example 2.3 in a permuted form. Since

$$\begin{bmatrix} G & M \\ -M & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} I & -\frac{1}{2}G \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ \frac{1}{2}G & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix},$$

the pencil Example 2.3 is equivalent to the Hamiltonian matrix

$$H = J \begin{bmatrix} I & \frac{1}{2}G \\ 0 & I \end{bmatrix} \begin{bmatrix} K & 0 \\ 0 & M^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -\frac{1}{2}G & I \end{bmatrix},$$

as is shown in [63]. There is no need to assemble the matrix  $H$ , nor is there any need to compute  $M^{-1}$  explicitly. One just needs to compute and use the Cholesky decomposition of  $M$  and one should also note that

$$H^{-1} = \begin{bmatrix} I & 0 \\ \frac{1}{2}G & 0 \end{bmatrix} \begin{bmatrix} K^{-1} & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} I & -\frac{1}{2}G \\ 0 & I \end{bmatrix} J^T,$$

is no less accessible than  $H$  itself. This is important, since if one wants the eigenvalues of  $H$  that are closest to the origin, one works with  $H^{-1}$ . If one wants the eigenvalues near a focal-point  $\tau$ , one might prefer to work with shift-and-invert, i.e.  $(H - \tau I)^{-1}$ . However, the shift destroys the Hamiltonian structure, so one needs ways to effect shifts while preserving the structure. One simple remedy is to work with the matrix

$$(H - \tau I)^{-1}(H + \tau I)^{-1},$$

which is not Hamiltonian but skew-Hamiltonian, (i.e. it satisfies  $(HJ)^T = -HJ$ .) If the problem is real and  $\tau$  is neither real nor purely imaginary, one works with the skew-Hamiltonian

$$(H - \tau I)^{-1}(H - \bar{\tau} I)^{-1}(H + \tau I)^{-1}(H + \bar{\tau} I)^{-1}$$

to stay within the real number system. Another possibility is to work with the Cayley-transform  $(H - \tau I)^{-1}(H + \tau I)$ , which is symplectic.

A Krylov subspace method will preserve Hamiltonian, skew-Hamiltonian, or symplectic structure if it generates vectors that span isotropic subspaces. The skew-Hamiltonian form is easiest to preserve, since Krylov subspaces generated by skew-Hamiltonian operators are automatically isotropic [63]. Consequently the standard Arnoldi method preserves the structure automatically in theory. In practice, however, the isotropy is steadily eroded by roundoff errors, so it must be enforced by an additional orthogonalization step. In the context of the Arnoldi process, this means that the vector  $q_{j+1}$  generated on step  $j$  must be made orthogonal to  $Jq_1, \dots, Jq_j$  as well as  $q_1, \dots, q_j$ . Furthermore, as in all practical Krylov subspace methods, in order to avoid that the storage capacity is exceeded, repeated implicit restarts in the spirit of Sorensen's implicitly restarted Arnoldi (IRA) process [53, 97] are needed. A method that includes all these features is the *skew-Hamiltonian implicitly-restarted Arnoldi algorithm (SHIRA)* of [63]. We will present some numerical results obtained with this method in Section 4.

### 3.4 Iterative projection methods for nonlinear problems

In this section we will discuss projection methods that work directly for general nonlinear eigenproblems. We have already pointed out that in this case the search spaces have to be expanded by directions that have a high approximation potential for the next desired eigenvector.

Assume that  $V$  is an orthonormal basis of the current search space. Let  $(\theta, y)$  be a solution of the projected problem  $V^H \mathcal{F}(\lambda) V y = 0$ , and let

$x = Vy$  be the corresponding Ritz vector. Then there are two candidates for expanding  $V$  suggested by the methods in Section 3.1:  $\hat{v} = x - \mathcal{F}(\sigma)^{-1}\mathcal{F}(\theta)x$  motivated by residual inverse iteration, and  $\tilde{v} = \mathcal{F}(\theta)^{-1}\mathcal{F}'(\theta)x$  corresponding to inverse iteration.

The following two subsections will take advantage of these directions. Expanding a given search space  $V$  by  $\hat{v}$  results in Arnoldi type methods treated in Subsection 3.4.1. Expanding it by the direction of inverse iteration  $v = \mathcal{F}(\theta)^{-1}\mathcal{F}'(\theta)x$  requires the solution of a large linear system in every iteration step. As in the case of linear eigenproblems this can be avoided by a Jacobi–Davidson approach considered in Subsection 3.4.2.

### 3.4.1 Arnoldi type methods

We consider the expansion of  $V$  by  $\hat{v} = x - \mathcal{F}(\sigma)^{-1}\mathcal{F}(\theta)x$ , where  $\sigma$  is a fixed shift (not too far away from the eigenvalue targeted at).

In Arnoldi-like methods the new search direction is orthonormalized against the previous ansatz vectors. Since the Ritz vector  $x$  is contained in the span of  $V$ , one may choose the new direction  $v = \mathcal{F}(\sigma)^{-1}\mathcal{F}(\theta)x$  as well. For the linear problem  $\mathcal{F}(\lambda) = A - \lambda B$  this is exactly the Cayley transformation with pole  $\sigma$  and zero  $\theta$ , and since

$$(A - \sigma B)^{-1}(A - \theta B) = I + (\theta - \sigma)(A - \sigma B)^{-1}B$$

and Krylov spaces are shift-invariant, the resulting projection method expanding  $V$  by  $v$  is nothing else but the shift-and-invert Arnoldi method.

If it is too expensive to solve the linear system  $\mathcal{F}(\sigma)v = \mathcal{F}(\theta)x$  for  $v$ , one may choose as new direction  $v = M\mathcal{F}(\theta)x$  with  $M \approx \mathcal{F}(\sigma)^{-1}$  which corresponds in the linear case to an inexact Cayley-transform or a preconditioned Arnoldi method. These variants have been introduced in [60] for quadratic eigenvalue problems and in [105, 109] for general nonlinear eigenproblems and are often called *nonlinear Arnoldi method* despite the fact that differently from the linear case no Krylov space is determined in the course of the algorithm.

Since the speed of convergence depends crucially on  $|\sigma - \lambda|$ , it will be advisable to change the shift or more generally the preconditioner  $M$  in the course of the algorithm if the convergence to the current eigenvalue becomes too slow. If one does this then the resulting method generalizes the *rational Krylov method* that was developed for linear problems in [83]. Thus the name *nonlinear rational Krylov method* would be appropriate as well. But this notation was already introduced in [84, 85] for a different method which

we will also discuss below. A template for the nonlinear Arnoldi approach is given by the following algorithm.

---

**Algorithm 5** Nonlinear Arnoldi Method

---

- 1: start with an initial shift  $\sigma$  and an initial basis  $V$ ,  $V^H V = I$ ;
  - 2: determine a preconditioner  $M \approx \mathcal{F}(\sigma)^{-1}$ ,
  - 3: **for**  $m = 1, 2, \dots$ , number of wanted eigenvalues **do**
  - 4:   compute appropriate eigenvalue  $\mu$  and corresponding eigenvector  $y$  of the projected problem  $\mathcal{F}_V(\mu)y := V^H \mathcal{F}(\mu)V y = 0$ .
  - 5:   determine Ritz vector  $u = Vy$  and residual  $r = \mathcal{F}(\mu)u$
  - 6:   **if**  $\|r\|/\|u\| < \epsilon$  **then**
  - 7:     accept approximate eigenpair  $\lambda_m = \mu$ ,  $x_m = u$ ,
  - 8:     **if**  $m ==$  number of desired eigenvalues **then** STOP **end if**
  - 9:     choose new shift  $\sigma$  and determine a preconditioner  $M \approx \mathcal{F}(\sigma)^{-1}$  if indicated
  - 10:    restart if necessary
  - 11:    choose approximations  $\mu$  and  $u$  to next eigenvalue and eigenvector
  - 12:    determine residual  $r = \mathcal{F}(\mu)u$
  - 13:    **end if**
  - 14:     $v = Mr$
  - 15:     $v = v - VV^H v$ ,  $\tilde{v} = v/\|v\|$ ,  $V = [V, \tilde{v}]$
  - 16:    reorthogonalize if necessary
  - 17: **end for**
- 

**Remark 3.1** There are many details that have to be considered when implementing a nonlinear Arnoldi method according to Algorithm 5. A detailed discussion is given in [106, 111], but the following general comments should be observed.

- (i) In Step 1 of Algorithm 5 any pre-information such as known approximate eigenvectors of problem (1) or eigenvectors of contiguous problems can and should be used. If no information on eigenvectors is at hand, and one is interested in eigenvalues close to a focal point  $\tau \in D$ , then one can choose an initial vector at random, execute a few Arnoldi steps for the linear eigenproblem  $\mathcal{F}(\tau)u = \theta u$  or  $\mathcal{F}(\tau)u = \theta \mathcal{F}'(\tau)u$ , and choose  $V$  by orthogonalizing eigenvectors corresponding to small eigenvalues in modulus. Starting with a random vector without this preprocessing usually does not lead to convergence.
- (ii) The preconditioner in Step 2 should be chosen on the basis of the

underlying problem. If this is not available, then full or incomplete sparse  $LU$  decompositions of  $\mathcal{F}(\sigma)$ , see [86], present an alternative.

- (iii) A crucial point in iterative methods for general nonlinear eigenvalue problems when approximating more than one eigenvalue is to inhibit the method to converge to the same eigenvalue repeatedly. For linear eigenvalue problems this is easy to do by using Schur forms or generalized Schur forms for the projected problem and then locking or purging certain eigenvalues. For nonlinear problems, however, such Schur forms do not exist and this presents one of the most difficult tasks in achieving good convergence. See [15, 23, 57, 60, 105] for different approaches in this direction.
- (iv) Since the residual inverse iteration with a fixed shift  $\sigma$  converges linearly, and the convergence rate satisfies  $\mathcal{O}(|\sigma - \lambda|)$ , it is reasonable to update the preconditioner if the convergence measured by the quotient of the last two residual norms before convergence has become too slow.

For several other recent variations and generalizations of the Arnoldi method for quadratic or general polynomial eigenvalue problems, see [7, 25, 36, 60, 100].

### 3.4.2 Jacobi–Davidson type methods

Arnoldi type methods are quite efficient in solving sparse nonlinear eigenproblems if an accurate preconditioner  $M \approx \mathcal{F}^{-1}$  is at hand. If this is not the case, then the convergence deteriorates considerably. In this situation Jacobi–Davidson type methods offer an appealing alternative.

A natural generalization of the Jacobi–Davidson method for linear eigenproblems which was already suggested in [92, 95] for polynomial eigenvalue problems and which was studied in [110] and [15] for general nonlinear eigenproblems is the following one: Suppose that the columns of  $V \subset \mathbb{C}^n$  form an orthonormal basis of the current search space, and let  $(x, \theta)$  be a Ritz pair of (1) with respect to  $V$ , i.e.  $V^H \mathcal{F}(\theta) V y = 0$ ,  $x = V y$ . Then we consider the correction equation

$$\left( I - \frac{p x^H}{x^H p} \right) \mathcal{F}(\theta) \left( I - \frac{x x^H}{x^H x} \right) z = -r, \quad z \perp x \quad (18)$$

where  $p := \mathcal{F}'(\theta)x$  and  $r := \mathcal{F}(\theta)x$ .

Equation (18) can be rewritten as  $\mathcal{F}(\theta)z - \alpha p = -r$ , where  $\alpha$  has to be chosen such that  $z \perp x$ . Solving for  $z$  we obtain

$$z = -x + \alpha \mathcal{F}(\theta)^{-1} p = -x + \alpha \mathcal{F}(\theta)^{-1} \mathcal{F}'(\theta)x,$$

and  $x = Vy$  yields that  $\tilde{z} := \mathcal{F}(\theta)^{-1} \mathcal{F}'(\theta)x \in \text{span}[V, z]$ .

Hence, as in the linear case the new search space  $\text{span}[V, z]$  contains the vector obtained by one step of inverse iteration with shift  $\theta$  and initial vector  $x$ , and again we may expect quadratic or even cubic convergence of the resulting iterative projection method, if the correction equation (18) is solved exactly.

As in the linear case the correction equation does not have to be solved exactly to maintain fast convergence, but usually a few steps of a Krylov solver with an appropriate preconditioner suffice to obtain a good expansion direction of the search space. The approximate solution of the correction equation has to be orthogonal to  $x$ . If (18) is solved by a Krylov solver and the initial approximation is orthogonal to  $x$  then all iterates are orthogonal to  $x$  as well.

In the correction equation (18), the operator  $\mathcal{F}(\theta)$  is restricted to map the subspace  $x^\perp$  to  $(\mathcal{F}'(\theta)x)^\perp$ . Hence, if  $K \approx \mathcal{F}(\theta)$  is a preconditioner of  $\mathcal{F}(\theta)$  then a preconditioner for an iterative solver of (18) should be modified correspondingly to

$$\tilde{K} := \left(I - \frac{px^H}{x^H p}\right) K \left(I - \frac{xx^H}{x^H x}\right).$$

With left-preconditioning, equation (18) becomes

$$\tilde{K}^{-1} \left(I - \frac{px^H}{x^H p}\right) \mathcal{F}(\theta) \left(I - \frac{xx^H}{x^H x}\right) z = -\tilde{K}^{-1} r, \quad z \perp x. \quad (19)$$

It was already demonstrated in [94] for linear problems that taking into account the projectors in the preconditioner, i.e. using  $\tilde{K}$  instead of  $K$  in a preconditioned Krylov solver, raises the cost only slightly. In every step one has to solve one linear system  $Kw = y$ , and to initialize the solver requires only one additional solve. The transfer to (19) is straightforward [15].

A template for the Jacobi–Davidson method for the nonlinear eigenvalue problem (1) is very similar to the one in Algorithm 5. We only have to replace Step 14 by:

Find an approximate solution of the correction equation

$$\left(I - \frac{\mathcal{F}'(\mu)uu^H}{u^H \mathcal{F}'(\mu)u}\right) \mathcal{F}(\mu) \left(I - \frac{uu^H}{u^H u}\right) t = -r \quad (20)$$



by a preconditioned Krylov solver.

Note that the update rule for the preconditioner used in the Arnoldi type method does not make sense for the Jacobi–Davidson algorithm because of the unpredictable convergence behavior of Krylov solvers. One way is to update the shift if the Krylov solver has not converged in a given number of iterations.

In [101] the efficiency of the Jacobi–Davidson method is demonstrated for large quadratic eigenproblems modeling the propagation of sound waves in a room with sound–absorbing walls. In [39] a gyroscopic eigenproblem  $(\lambda^2 M + \lambda(G + \varepsilon D) + K)x = 0$  is considered, where  $K = K^T$ ,  $M = M^T$ ,  $G = -G^T$ , and  $\varepsilon D$  represents the damping of the system. Since the damping is assumed to be small it is suggested to determine eigenpairs  $(\lambda_j, x_j)$  of the conservative gyroscopic problem  $(\lambda^2 M + \lambda G + K)x = 0$  in the wanted region by the structure preserving method SHIRA from Section 3.3, and to improve these approximations by the Jacobi–Davidson method for the original quadratic problem.

### 3.4.3 Rational Krylov method

A different approach was developed in [81, 82, 84, 85], which generalizes the rational Krylov method for linear eigenproblems [83] to sparse nonlinear eigenvalue problems by nesting the linearization of problem (1) (by regula falsi) and the solution of the resulting linear eigenproblem (by Arnoldi’s method). Hence, in this approach the two numerical subtasks, i.e. reducing the large dimension to a much smaller one and solving a nonlinear eigenproblem are attacked simultaneously. This method was applied in [31, 32] to the rational eigenvalue problem (9) governing damped vibrations of a structure.

Linearizing the nonlinear family  $\mathcal{F}(\lambda)$  by Lagrange interpolation between two points  $\mu$  and  $\sigma$  one gets

$$\mathcal{F}(\lambda) = \frac{\lambda - \mu}{\sigma - \mu} \mathcal{F}(\sigma) + \frac{\lambda - \sigma}{\mu - \sigma} \mathcal{F}(\mu) + \text{higher order terms.}$$

Keeping  $\sigma$  fixed for several steps, iterating on  $\mu$ , neglecting the remainder in the Lagrange interpolation, and multiplying by  $\mathcal{F}(\sigma)^{-1}$  from the left one obtains

$$\mathcal{F}(\sigma)^{-1} \mathcal{F}(\lambda_{j-1}) w = \theta w \quad \text{with} \quad \theta = \frac{\lambda_j - \lambda_{j-1}}{\lambda_j - \sigma}, \quad (21)$$

predicting a singularity at

$$\lambda_j = \lambda_{j-1} + \frac{\theta}{1-\theta}(\lambda_{j-1} - \sigma).$$

For large and sparse matrices the linearization (21) is combined with a linear Arnoldi process. After  $j$  steps, approximations  $\lambda_1, \dots, \lambda_j$  to an eigenvalue, orthonormal vectors  $v_1, \dots, v_j$ , and an upper Hessenberg matrix  $H_{j,j-1} \in \mathbb{C}^{j,j-1}$  are generated, such that the Arnoldi recursion

$$\mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda_{j-1})V_{j-1} = V_j H_{j,j-1}$$

is fulfilled (at least approximately), where  $V_j = [v_1, \dots, v_j]$ .

Updating the matrix  $H_{j,j-1}$  according to the linear theory yields

$$\tilde{H}_{j+1,j} = \begin{bmatrix} H_{j,j-1} & k_j \\ 0 & \|r_\perp\| \end{bmatrix},$$

where  $k_j = V_j^H r_j$ ,  $r_j = \mathcal{F}(\lambda_j)v_j$ , and  $r_\perp = r_j - V_j V_j^H v_j$ , which due to the nonlinearity of  $\mathcal{F}(\cdot)$  violates the next Arnoldi relation

$$\mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda_j)V_j = V_{j+1}\tilde{H}_{j+1,j}, \quad v_{j+1} = v_\perp / \|v_\perp\|.$$

To satisfy it at least approximately one takes advantage of Lagrangian interpolation

$$\mathcal{G}(\lambda_j) \approx \frac{\lambda_j - \sigma}{\lambda_{j-1} - \sigma} \mathcal{G}(\lambda_{j-1}) - \frac{\lambda_j - \lambda_{j-1}}{\lambda_{j-1} - \sigma} I = \frac{1}{1-\theta} \mathcal{G}(\lambda_{j-1}) - \frac{\theta}{1-\theta} I,$$

where  $\mathcal{G}(\lambda) := \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)$ , and updates  $H$  according to

$$H_{j+1,j} = \begin{bmatrix} \frac{1}{1-\theta} H_{j,j-1} - \frac{\theta}{1-\theta} I_{j,j-1} & k_j \\ 0 & \|r_\perp\| \end{bmatrix}.$$

This gives a first version of the rational Krylov method, which unfortunately is not very efficient. In [84] it was thus suggested to modify  $\lambda$  and  $H$  in an inner iteration until the residual  $r = \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)V_j s$  is enforced to be orthogonal to  $V_j$ , and to expand the search space only after the inner iteration has converged which yields Algorithm 6.

The inner iteration is nothing else but a solver of the projected problem

$$V_j^H \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)V_j s = 0. \quad (22)$$

---

**Algorithm 6** Rational Krylov method

---

- 1: start with initial vector  $V = [v_1]$  with  $\|v_1\| = 1$ , initial  $\lambda$  and  $\sigma$ ; set  $j = 1$
  - 2: set  $h_j = 0_j$ ;  $s = e_j$ ;  $x = v_j$ ;
  - 3: compute  $r = \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)x$  and  $k_j = V_j^H r$
  - 4: **while**  $\|k_j\| > \text{ResTol}$  **do**
  - 5:   orthogonalize  $r = r - V_j^H k_j$
  - 6:   set  $h_j = h_j + k_j s_j^{-1}$
  - 7:    $\theta = \min \text{eig } H_{j,j}$  with corresponding eigenvector  $s$
  - 8:    $x = V_j s$
  - 9:   update  $\lambda = \lambda + \frac{\theta}{1-\theta}(\lambda - \sigma)$
  - 10:   update  $H_{j,j} = \frac{1}{1-\theta}H_{j,j} - \frac{1}{1-\theta}I$
  - 11:   compute  $r = \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)x$  and  $k_j = V_j^H r$
  - 12: **end while**
  - 13: compute  $h_{j+1,j} = \|r\|$
  - 14: **if**  $|h_{j+1,j}s_j| > \text{EigTol}$  **then**
  - 15:    $v_{j+1} = r/h_{j+1,j}$ ;  $j = j + 1$ ; GOTO 2:
  - 16: **end if**
  - 17: Accept eigenvalue  $\lambda_i = \lambda$  and eigenvector  $x_i = x$
  - 18: If more eigenvalues wanted, choose next  $\theta$  and  $s$ , and GOTO 8:
- 

---

**Algorithm 7** Rational Krylov method, an iterative projection method

---

- 1: start with initial vector  $V = [v_1]$  with  $\|v_1\| = 1$ , initial  $\lambda$  and  $\sigma$
  - 2: **for**  $j = 1, 2, \dots$  until convergence **do**
  - 3:   solve projected eigenproblem  $V^H \mathcal{F}(\sigma)^{-1} \mathcal{F}(\lambda) V s = 0$  for  $(\lambda, s)$   
by inner iteration
  - 4:   compute Ritz vector  $x = V s$  and residual  $r = \mathcal{F}(\sigma)^{-1} \mathcal{F}(\lambda) x$
  - 5:   orthogonalize  $r = r - V V^H r$
  - 6:   expand searchspace  $V = [V, r/\|r\|]$
  - 7: **end for**
-

Hence, although motivated originally in a different way, the rational Krylov method is an iterative projection method, where the nonlinear eigenproblem  $\mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)x = 0$  is projected to a search space  $V$ , and  $V$  is expanded by the orthogonal complement (with respect to  $V$ ) of the residual  $r = \mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)V_s$  of the Ritz pair and one ends up with Algorithm 7.

The inner iteration in Step 3 of Algorithm 7 can be replaced by any dense solver of Section 3.1 and numerical examples in [43] demonstrate that the method can be accelerated considerably this way. It is a disadvantage of the rational Krylov method that symmetry properties which the original problem may have are destroyed if the projected problem (22) is considered instead of  $V_j^H \mathcal{F}(\lambda) V_j s = 0$  in the Arnoldi method or the Jacobi–Davidson algorithm. But on the other hand, the solvers in Section 3.1 need the explicit form of the projected problem whereas the inner iteration in Algorithm 6 only needs a procedure that yields the vector  $\mathcal{F}(\sigma)^{-1}\mathcal{F}(\lambda)x$  for a given  $x$ .

Iterative projection methods for nonlinear eigenvalue problems have been studied extensively in recent years, but still have not reached a stage of maturity as have the methods for linear problems. *More research is needed to improve the current methods, to analyze their stability and convergence behavior, in particular in finite precision, and to generate implementations that can be used in a simple and robust way by non-experts.*

## 4 Numerical examples

To demonstrate the iterative projection methods from the previous Sections, we give several numerical examples that have already been mentioned in the introduction.

### 4.1 Vibrations of a fluid–solid structure

Free vibrations of a tube bundle immersed in a slightly compressible fluid are governed under certain conditions by a boundary value problem with nonlocal boundary conditions depending rationally on the eigenparameter, see [17, 18, 73]. Discretizing by finite elements one obtains a rational matrix eigenvalue problem

$$\mathcal{F}(\lambda)x := -Kx + \lambda Mx + \sum_{j=1}^k \frac{\lambda}{\sigma_j - \lambda} C_j x = 0,$$

where  $K$ ,  $M$ , and  $C_j$  are symmetric matrices,  $K$  and  $C_j$  are positive semidefinite,  $M$  is positive definite, and  $0 =: \sigma_0 < \sigma_1 < \dots < \sigma_k < \sigma_{k+1} := \infty$

are given positive numbers. Then, in each of the intervals  $(\sigma_j, \sigma_{j+1})$ ,  $j = 0, \dots, k$ , this problem satisfies the conditions of the min-max characterization (16).

We consider a problems of dimension  $n = 22654$  with one pole  $\sigma_1 = 1$  which has 11 eigenvalues  $\lambda_1 \leq \dots \leq \lambda_{11}$  in the interval  $J_1 = (0, 1)$ , see [59], and a large number of eigenvalues greater than 1, 10 of which are contained in the interval  $(1, 4)$ .

We determined approximations to the eigenvalues in  $[0, 1)$  by the Arnoldi method, the Jacobi–Davidson method, where in both cases the projected nonlinear eigenproblems were solved by safeguarded iteration, and by the rational Krylov method, where the projected rational eigenproblems were solved by linearizing the equivalent quadratic eigenproblem

$$(1 - \lambda)V^T \mathcal{F}(\sigma)^{-1} \mathcal{F}(\lambda)Vy = 0.$$

All three methods were able to find all 11 eigenvalues. The original rational Krylov method (Algorithm 6) as implemented in [42] turned out to depend very sensitively on the initial pole  $\sigma$  and the initial approximation to an eigenvalue, and was able to find at most 8 eigenvalues in the interval  $[0, 1)$ .

Enumerating the eigenvalues according to Section 3.1 the smallest eigenvalue in the interval  $(1, 4)$  turns out to be a 6-th eigenvalue. Starting the Arnoldi method and the Jacobi–Davidson method with the invariant subspace of the linear eigenproblem, see [106],

$$\left(K + \frac{\lambda}{\lambda - 1}C_1\right)x = \mu Mx, \quad \lambda = 1 + \varepsilon, \quad \varepsilon > 0,$$

corresponding to the 6 largest eigenvalues, both methods were able to find all eigenvalues in the interval  $(1, 4)$ . The rational Krylov method destroys the symmetry of the problem, and the enumeration of Subsection 3.1 does not apply. None of the implementations of the rational Krylov methods was able to find more than one or two eigenvalues in the interval  $(1, 4)$ .

The numerical experiments were run under MATLAB 6.5 on an Intel Centrino M processor with 1.7 GHz and 1 GB RAM. Table 1 summarizes the properties of the iterative projection methods under consideration, where `Iter.` is the total number of iteration steps, `LU fact.` is the number of LU factorizations, `CPU` denotes the CPU time in seconds, and `nlin.sol.` the CPU time required for solving the projected eigenvalue problems. In either case only a very small share of the total CPU time was needed for solving the projected problems.

Method	Iter.	LU fact.	CPU [s]	nlin.sol. [s]
Arnoldi	34	2	14.93	0.13
Jacobi–Davidson	37	3	112.84	0.15
rational Krylov	40	2	70.80	0.22
interval (1, 4)				
Arnoldi	36	2	17.35	0.17
Jacobi–Davidson	37	5	125.87	0.23

Table 4.1 Iterations and cpu-times for fluid–solid structure problem

## 4.2 Damped vibrations of a structure

As second example we consider the free vibrations of a finite element model as in (9) of dimension 10704 for a solid using a viscoelastic constitutive relation to describe the behavior of the material.

For symmetry reasons we determined only eigenvalues with negative imaginary part, and we computed 30 of them one after another with decreasing imaginary part. The nonlinear projected eigenproblems (after multiplying by  $1 + b\omega$ ) were solved by linearization, and the iteration was terminated if the norm of the residual was less than  $10^{-6}$ .

The Arnoldi method without restarts needed 144 iteration steps, and a CPU time of 707.0 seconds to find all 30 eigenvalues with maximal negative imaginary part (i.e. the average number of iteration steps to determine an eigenvalue is less than 5). No update of the preconditioner was necessary.

The dominant share of the CPU time, namely 469.9 seconds was consumed by the solver of the projected nonlinear eigenproblems. Figure 2 on the left displays the development of the time consumption of the entire iteration and the share of the nonlinear eigensolver. x-marks indicate found eigenvalues. It demonstrates the necessity of restarts, since the superlinear time consumption is mainly caused by the eigensolver.

We restarted the Arnoldi process if the dimension of the search space exceeded 50 with an orthogonal basis of the space spanned by the already determined eigenvectors. The method needed 3 restarts, and again all 30 eigenvalues with maximal negative imaginary part were found by the Arnoldi method requiring 139 iterations, and 4 updates of the preconditioner. The total CPU time was 199.6 seconds. Solving the projected eigenproblems and updating the preconditioners required 25.0 and 30.8 seconds, respectively. The right picture in Figure 2 demonstrates the time consumption for this experiment. Immediately after a restart the speed of convergence is slowed down. On the other hand this delay yields an update of the preconditioner

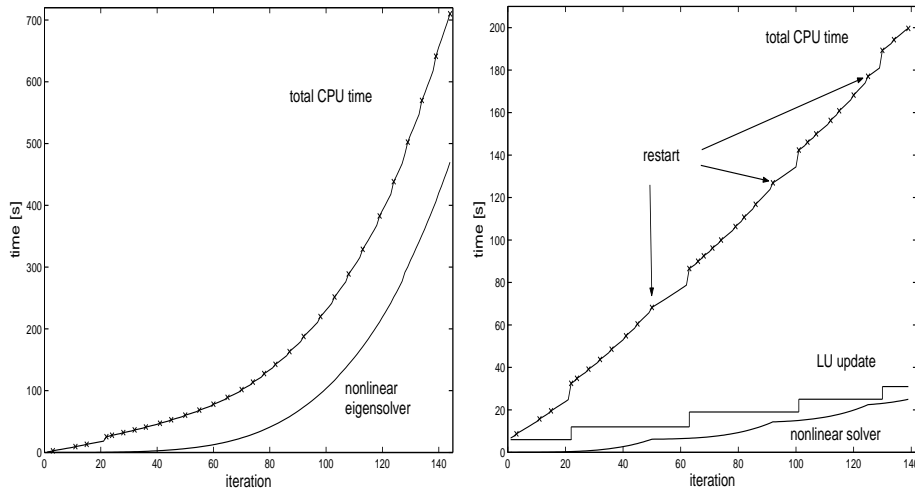


Figure 2: Arnoldi method without and with restarts

accelerating the convergence, such that the total number of iteration steps is reduced from 144 to 139.

The Jacobi–Davidson method and the rational Krylov method according to Algorithm 7 are much slower than the Arnoldi method, but they show similar behavior. Without restart a substantial share of the total CPU time is consumed by the solver of the projected nonlinear eigenproblems, and both methods can be accelerated by restarts. Details about these methods are contained in Table 4.2. The rational Krylov method with inner iteration converges for this problem, although it is very slow. To determine the 30 wanted eigenvalues, 366 iterations are necessary requiring 2248.5 seconds.

Method	Iter.	LU fact.	CPU [s]	nlin.sol. [s]
Arnoldi	144	2	707.0	469.9
Arnoldi, restarted	139	5	199.6	25.0
Jacobi–Davidson	111	9	1050.5	161.2
Jacobi–Davidson, restarted	109	12	914.4	18.9
rational Krylov	147	3	1107.1	465.3
rational Krylov, restarted	147	4	647.8	28.5

Table 4.2: Iterations and cpu times for damped vibrations of a structure

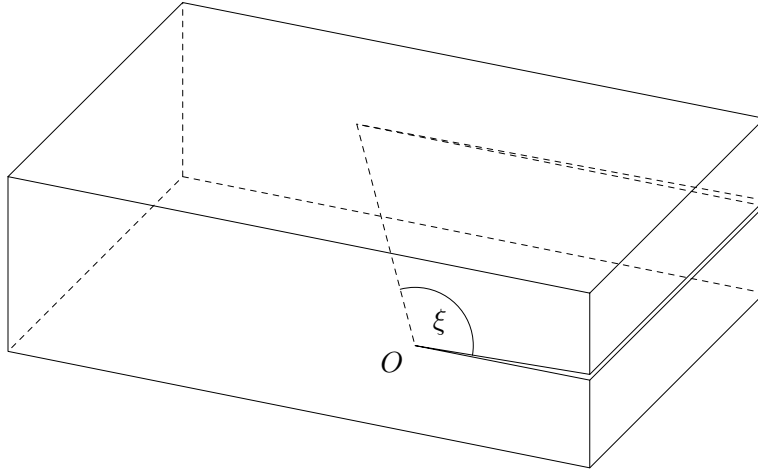


Figure 3: Illustration of the crack example

### 4.3 Singularity exponents

For a numerical example demonstrating the structure preserving methods, we consider the following problem which originates from [8, 51], for details see [4]. Consider a specimen in form of a brick with a crack as illustrated in Figure 3. The (homogeneous) material has Poisson ratio  $\nu = 0.32$ , the Young modulus does not influence the result and can be set arbitrarily. The stress concentration near the point  $O$  where the crack intersects the surface of the brick can be investigated within the linear elasticity framework, as long as the material is brittle, see e.g. [50, 51, 52]. The stress field is derived from the displacement field which can be represented by a regular part and several singular terms of the form  $kr^\alpha u(\varphi, \theta)$  where  $(r, \varphi, \theta)$  are spherical coordinates centered in the point of interest,  $k$  is called stress intensity factor, and  $\alpha$  is the characteristic (singular) exponent with the associated mode  $u$ . These terms are singular when  $\Re\alpha < 1$  and  $\Re\alpha \neq 0$ .

Mathematically, the pair  $(\alpha, u)$  is an eigenpair of a quadratic operator pencil [44, 51]. A finite element approximation [5] leads to a finite-dimensional quadratic eigenvalue problem. After substitution  $\lambda = \alpha + 0.5$  this problem has the more convenient form (3). Figure 4 displays the real part of the (approximated) eigenvalues from the strip  $0 < \Re\alpha < 1$  for the whole possible range of the angle  $\xi$  between the crack and the surface. Which angle actually appears depends on the applied load. In [4] several structure preserving linearizations and corresponding iterative projection



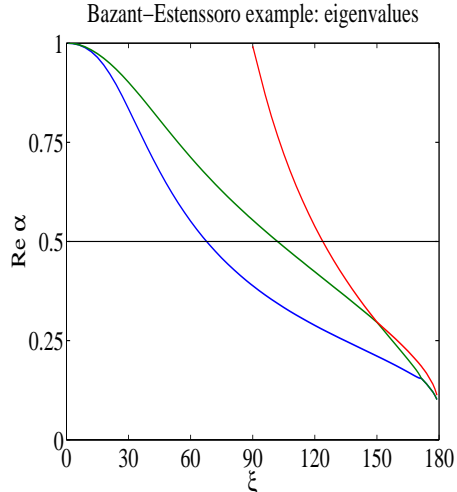


Figure 4: Real part of the interesting eigenvalues against the angle  $\xi$ .

methods (in a FORTRAN 77 implementation) were compared for this by computing 9 eigenvalues in order to obtain approximations to the 2 – 3 eigenvalues nearest to the imaginary axis. A comparison of the Hamiltonian implicitly restarted Lanczos method HIRL of [114] (with focal point  $\tau = 0.0$ ) and the skew-Hamiltonian implicitly-restarted Arnoldi algorithm (SHIRA) of [63] with focal points  $\tau = 1.0$  and  $\tau = 0.0$ , as well as a simple non-structured method is given in Figure 5 for the angle  $\xi = 120^\circ$ , where the three eigenvalues of interest are well separated. The total CPU time is presented for various discretization parameters  $N$ , which is the size of the coefficient matrices  $K$ ,  $G$ , and  $M$ .

The comparison shows that despite a higher complexity per iteration step, the structure preserving method SHIRA proves to be a competitive algorithm to compute eigenvalues of even or odd quadratic pencils, in particular if the focal point is not close to an eigenvalue. For more details see [4].

## 5 Conclusion

We have discussed the numerical solution of large scale nonlinear eigenvalue problems. The development of efficient and reliable numerical methods for these classes of problems remains a major challenge. More research is needed, in particular in the direction of error and condition estimation as

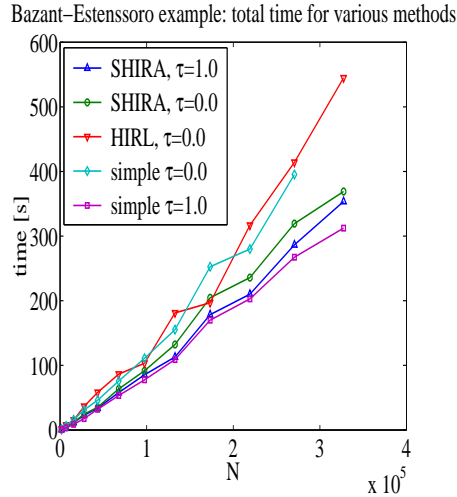


Figure 5: CPU times for crack with  $\xi = 120^\circ$  and various  $N$ .

well as the implementation of appropriate software packages such as those that are available for linear eigenvalue problems.

## References

- [1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, PA, second edition, 1995.
- [2] P.M. Anselone and L.B. Rall. The solution of characteristic value-vector problems by Newton's method. *Numer. Math.*, 11:38–45, 1968.
- [3] T. Apel, V. Mehrmann, and D. Watkins. Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures. *Comp. Meth. Appl. Mech. Eng.*, 191:4459–4473, 2002.
- [4] T. Apel, V. Mehrmann, and D. Watkins. Numerical solution of large scale structured polynomial eigenvalue problems. In *Foundations of Computational Mathematics*. Springer Verlag, 2003.
- [5] T. Apel, A.-M. Sändig, and S. I. Solov'ev. Computation of 3D vertex singularities for linear elasticity: Error estimates for a finite element

- method on graded meshes. *Math. Modeling Numer. Anal. (M2AN)*, 2004. To appear.
- [6] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the Solution of Algebraic Eigenvalue Problems*. SIAM Publications, Philadelphia, 2000.
- [7] Z. Bai and Y. Su. A second-order Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, To appear, 2005.
- [8] Z. P. Bažant and L.F. Estenssoro. Surface singularity and crack propagation. *Internat. J. Solids Structures*, 15:405–426, 1979.
- [9] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of skew Hamiltonian/Hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 24:165–190, 2002.
- [10] P. Benner and H. Faßbender. An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra Appl.*, 263:75–111, 1997.
- [11] P. Benner and H. Faßbender. An implicitly restarted symplectic Lanczos method for the symplectic eigenvalue problem. *Berichte aus der Technomathematik*, Report 98–01, FB3 – Mathematik und Informatik, Universität Bremen, 28334 Bremen, FRG, 1998.
- [12] P. Benner and H. Faßbender. The symplectic eigenvalue problem, the butterfly form, the *SR* algorithm, and the Lanczos method. *Linear Algebra Appl.*, 275/276:19–47, 1998.
- [13] P. Benner, V. Mehrmann, V. Sima, S. Van Huffel, and A. Varga. SLICOT - a subroutine library in systems and control theory. *Appl. Comput. Contr., Sign., Circ.*, 1:499–532, 1999.
- [14] P. Benner, E.S. Quintana-Ortí, and G. Quintana-Ortí. Parallel model reduction of large-scale linear descriptor systems via balanced truncation. In *High Performance Computing for Computational Science. Proceedings of the 6th International Meeting VECPAR'04*, pages 65–78. Valencia, Spain, 2004.
- [15] T. Betcke and H. Voss. A Jacobi–Davidson–type projection method for nonlinear eigenvalue problems. *Future Generation Computer Systems*, 20(3):363 – 372, 2004.

- [16] L.S. Blackford, J. Choi, A. Cleary, E.D. Azevedo, J. Demmel, I. Dhillon, J. Dongarra, G. Henry, A. Petit, K. Stanley, D. Walker, and R. Whaley. *ScaLAPACK Users' Guide*. SIAM, Philadelphia, PA, second edition, 1997.
- [17] C. Conca, J. Planchard, and M. Vanninathan. Existence and location of eigenvalues for fluid-solid structures. *Comput. Meth. Appl. Mech. Engrg.*, 77:253 – 291, 1989.
- [18] C. Conca, J. Planchard, and M. Vanninathan. *Fluid and Periodic Structures*, volume 24 of *Research in Applied Mathematics*. Masson, Paris, 1995.
- [19] S.H. Crandall. Iterative procedures related to relaxation methods for eigenvalue problems. *Proc. Royal Soc. London*, 207:416–423, 1951.
- [20] R.J. Duffin. A minmax theory for overdamped networks. *J.Rat.Mech.Anal.*, 4:221 – 233, 1955.
- [21] R.J. Duffin. The rayleigh-ritz method for dissipative and gyroscopic systems. *Quart. Appl. Math.*, 18:215–221, 1960.
- [22] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Systems*. B. G. Teubner Stuttgart, 1998.
- [23] D.R. Fokkema, G.L.G. Sleijpen, and H.A. van der Vorst. Jacobi-Davidson style QR and QZ algorithms for the partial reduction of matrix pencils. *SIAM J. Sci. Comput.*, 20:94 – 125, 1998.
- [24] G. Freiling, V. Mehrmann, and H. Xu. Existence, uniqueness and parametrization of lagrangian invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 23:1045–1069, 2002.
- [25] R. Freund. Pade-type model reduction of second-order and higher-order linear dynamical systems. Technical report, Dept. of Mathematics, UC. Davis, Davis, Ca., 2004. <http://arxiv.org/abs/math/0410195>.
- [26] R. W. Freund. Krylov-subspace methods for reduced-order modeling in circuit simulation. *J. Comput. Appl. Math.*, 123:395–421, 2000.
- [27] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.
- [28] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.

- [29] K. P. Hadeler. Variationsprinzipien bei nichtlinearen Eigenwertaufgaben. *Arch. Rat. Mech. Anal.*, 30:297 – 307, 1968.
- [30] K. P. Hadeler. Nonlinear eigenvalue problems. In R. Ansorge, L. Colatz, G. Hämmerlin, and W. Tornig, editors, *Numerische Behandlung von Differentialgleichungen*, ISNM 27, pages 111–129. Birkhäuser, Stuttgart, 1975.
- [31] P. Hager. *Eigenfrequency Analysis. FE-Adaptivity and a Nonlinear Eigenvalue Problem*. PhD thesis, Chalmers University of Technology, Göteborg, 2001.
- [32] P. Hager and N.E. Wiberg. The rational Krylov algorithm for nonlinear eigenvalue problems. In B.H.V. Topping, editor, *Computational Mechanics for the Twenty-First Century*, pages 379 – 402. Saxe-Coburg Publications, Edinburgh, 2000.
- [33] N. Higham, D.S. Mackey, N. Mackey, and F. Tisseur. Linearizations of symmetric polynomials by symmetric pencils. Technical report, Univ. of Manchester, 2004. In preparation.
- [34] A. Hilliges. Numerische lösung von quadratischen eigenwertproblemen mit anwendungen in der schienenodynamik. Master’s thesis, TU Berlin, Berlin, FRG, 2004. Diplomarbeit, in German.
- [35] A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS)*. Jyväskylä, Finland, 2004. CD-ROM.
- [36] U.B. Holz. *Subspace Approximation Methods for Perturbed Quadratic Eigenvalue Problems*. PhD thesis, Department of Mathematics, Stanford University, 2002.
- [37] H.Y. Hu, E.H. Dowell, and L.N. Virgin. Stability estimation of high dimensional vibrating systems under state delay feedback control. *J. Sound and Vibration*, 214:497 – 511, 1998.
- [38] H.Y. Hu and Z.H. Wang. *Dynamics of Controlled Mechanical Systems with Delayed Feedback*. Springer, Berlin, 2002.
- [39] T.-M. Hwang, W.-W. Lin, and V. Mehrmann. Numerical solution of quadratic eigenvalue problems with structure-preserving methods. *SIAM J. Sci. Comput.*, 24:1283–1302, 2003.

- [40] N. K. Jain and K. Singhal. On Kublanovskaya's approach to the solution of the generalized latent value problem for functional  $\lambda$ -matrices. *SIAM J. Numer. Anal.*, 20:1062–1070, 1983.
- [41] N. K. Jain, K. Singhal, and K. Huseyin. On roots of functional lambda matrices. *Comput. Meth. Appl. Mech. Engrg.*, 40:277–292, 1983.
- [42] E. Jarlebring. Krylov methods for nonlinear eigenvalue problems. Master's thesis, Royal Institute of Technology, Dept.Numer.Anal.Comput.Sci., Stockholm, 2003.
- [43] E. Jarlebring and H. Voss. Rational Krylov for nonlinear eigenproblems, an iterative projection method. Technical Report 69, Section of Mathematics, Hamburg University of Technology, 2003. Submitted to *Appl.Math.*
- [44] V.A. Kozlov, V.G. Maz'ya, and J. Roßmann. Spectral properties of operator pencils generated by elliptic boundary value problems for the Lamé system. *Rostocker Math. Kolloq.*, 51:5–24, 1997.
- [45] V.N. Kublanovskaya. On an application of Newton's method to the determination of eigenvalues of  $\lambda$ -matrices. *Dokl. Akad. Nauk. SSR*, 188:1240 – 1241, 1969.
- [46] V.N. Kublanovskaya. On an approach to the solution of the generalized latent value problem for  $\lambda$ -matrices. *SIAM. J. Numer. Anal.*, 7:532 – 537, 1970.
- [47] P. Lancaster. Strongly stable gyroscopic systems. *Electr. J. Linear Algebra*, 5:53–66, 1999.
- [48] P. Lancaster. *Lambda-matrices and Vibrating Systems*. Dover Publications, Mineola, New York, second edition, 2002.
- [49] P. Lancaster and P. Psarrakos. A note on weak and strong linearizations of regular matrix polynomials. Unpublished 5-page note, 2004.
- [50] B. Lawn. *Fracture of Brittle Solids*. Cambridge University Press, 1993.
- [51] D. Leguillon. Computation of 3d-singularities in elasticity. In M. Costabel et al., editor, *Boundary value problems and integral equations in nonsmooth domains*, volume 167 of *Lect. Notes Pure Appl. Math.*, pages 161–170, New York, 1995. Marcel Dekker. Proceedings of the conference, held at the CIRM, Luminy, France, May 3-7, 1993.

- [52] D. Leguillon and E. Sanchez-Palencia. On 3D cracks intersecting a free surface in laminated composites. *Int. J. Fracture*, 99:25–40, 1999.
- [53] R. B. Lehoucq, D. C. Sorensen, and C. Yang. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. SIAM, Philadelphia, 1998.
- [54] J.-R. Li. *Model Reduction of Large Linear Systems via Low Rank System Gramians*. Ph.d. thesis, Massachusetts Institute of Technology, Department of Mathematics, 2000.
- [55] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Palindromic polynomial eigenvalue problems: Good vibrations from good linearizations. Technical report, TU Berlin, Inst. f. Mathematik, Berlin, Germany, 2004. In preparation.
- [56] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. Technical report, TU Berlin, Inst. f. Mathematik, Berlin, Germany, 2004. In preparation.
- [57] M. Markiewicz and H. Voss. A local restart procedure for iterative projection methods for nonlinear symmetric eigenproblems. Technical Report 82, Section of Mathematics, Hamburg University of Technology, 2004. To appear in Proc. of Algoritmy 2005.
- [58] MATLAB, Version 7.0. The MathWorks, inc., 24 Prime Park Way, Natick, MA 01760-1500, USA, 1996.
- [59] L. Mazurenko and H. Voss. On the number of eigenvalues of a rational eigenproblem. Technical Report 61, Section of Mathematics, Hamburg University of Technology, 2003.
- [60] K. Meerbergen. Locking and restarting quadratic eigenvalue solvers. *SIAM J. Sci. Comput.*, 22:1814 – 1839, 2001.
- [61] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.
- [62] V. Mehrmann and C. Shi. Analysis of higher order linear differential-algebraic systems. Preprint 2004/17, Institut für Mathematik, TU Berlin, D-10623 Berlin, FRG, 2004. url: <http://www.math.tu-berlin.de/preprints/>.

- [63] V. Mehrmann and D. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-hamiltonian/hamiltonian pencils. *SIAM J. Sci. Comput.*, 22:1905–1925, 2001.
- [64] V. Mehrmann and D. Watkins. Polynomial eigenvalue problems with hamiltonian structure. *Electr. Trans. Num. Anal.*, 13:106–113, 2002.
- [65] MODELICA – A unified object-oriented language for physical systems modeling. Model specification. Version 2.1. Modelica Association, 2004. Available from <http://www.modelica.org/documents/ModelicaSpec21.pdf>.
- [66] A. Neumaier. Residual inverse iteration for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 22:914 – 923, 1985.
- [67] M. R. Osborne. Inverse iteration, Newton’s method and nonlinear eigenvalue problems. In *The Contribution of Dr. J.H. Wilkinson to Numerical Analysis*, Symp. Proc. Series 19, pages 21–53. Inst. Math. Appl., London, 1978.
- [68] M.R. Osborne. A new method for the solution of eigenvalue problems. *Comput. J.*, 7:228–232, 1964.
- [69] M.R. Osborne and S. Michelson. The numerical solution of eigenvalue problems in which the eigenvalue parameter appears nonlinearly, with an application to differential equations. *Comput. J.*, 7:66–71, 1964.
- [70] A.M. Ostrowski. On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors i. *Arch. Rational Mech. Anal.*, 1:233–241, 1958.
- [71] T. Penzl. A cyclic low rank Smith method for large, sparse Lyapunov equations with applications in model reduction and optimal control. *SIAM J. Sci. Comput.*, 21:1401–1418, 2000.
- [72] N. Petersmann. *Substrukturtechnik und Kondensation bei der Schwingungsanalyse*, volume 76 of *Fortschrittberichte VDI, Reihe 11: Schwingungstechnik*. VDI Verlag, Düsseldorf, 1986.
- [73] J. Planchard. Eigenfrequencies of a tube bundle placed in a confined fluid. *Comput. Meth. Appl. Mech. Engrg.*, 30:75 – 93, 1982.
- [74] J.S. Przemieniecki. *Theory of Matrix Structural Analysis*. McGraw-Hill, New York, 1968.



- [75] A.C.M. Ran and L. Rodman. Stability of invariant Lagrangian subspaces I. In I. Gohberg, editor, *Operator Theory: Advances and Applications*, volume 32, pages 181–218. Birkhäuser-Verlag, Basel, Switzerland, 1988.
- [76] A.C.M. Ran and L. Rodman. Stability of invariant Lagrangian subspaces II. In H.Dym, S. Goldberg, M.A. Kaashoek, and P. Lancaster, editors, *Operator Theory: Advances and Applications*, volume 40, pages 391–425. Birkhäuser-Verlag, Basel, Switzerland, 1989.
- [77] E.H. Rogers. A minmax theory for overdamped systems. *Arch.Rat.Mech.Anal.*, 16:89 – 96, 1964.
- [78] K. Rothe. *Lösungsverfahren für nichtlineare Matrixeigenwertaufgaben mit Anwenden auf die Ausgleichelementmethode*. Verlag an der Lotterbek, Ammersbek, 1989.
- [79] K. Rothe. Least squares element method for boundary eigenvalue problems. *Internat. J. Numer. Methods Engrg.*, 33:2129–2143, 1992.
- [80] A. Ruhe. Algorithms for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 10:674 – 689, 1973.
- [81] A. Ruhe. Computing nonlinear eigenvalues with spectral transformation Arnoldi. *ZAMM*, 76:S2: 17 – 20, 1996.
- [82] A. Ruhe. Rational Krylov for nonlinear eigenvalues, 1996. Talk presented at the Householder XIII Conference at Pontresina, Switzerland.
- [83] A. Ruhe. Rational Krylov, a practical algorithm for large sparse non-symmetric matrix pencils. *SIAM J. Sci. Comp.*, 19(5):1535–1551, 1998.
- [84] A. Ruhe. A rational Krylov algorithm for nonlinear matrix eigenvalue problems. *Zapiski Nauchnyh Seminarov POMI*, 268:176 – 180, 2000.
- [85] A. Ruhe. Rational Krylov for large nonlinear eigenproblems. Technical report, Royal Institute of Technology, Stockholm, Sweden, 2004. To appear in Proceedings of PARA'04, Lyngby, Denmark, 2004.
- [86] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM Publications, Philadelphia, second edition, 2003.
- [87] W. Schiehlen. *Advanced multibody system dynamics*. Kluwer, Stuttgart, 1993.

- [88] F. Schmidt, T. Friese, L. Zschiedrich, and P. Deuffhard. Adaptive multigrid methods for the vectorial Maxwell eigenvalue problem for optical waveguide design. In W. Jäger and H.-J. Krebs, editors, *Mathematics. Key Technology for the Future*, pages 279–292. Springer-Verlag, 2003.
- [89] H. Schmitz, K. Volk, and W. L. Wendland. On three-dimensional singularities of elastic fields near vertices. *Numer. Methods Partial Differ. Equations*, 9:323–337, 1993.
- [90] H.R. Schwarz. *Methode der finiten Elemente*. Teubner, Stuttgart, 1984.
- [91] C. Shi. *Linear Differential-Algebraic Equations of Higher-Order and the Regularity or Singularity of Matrix Polynomials*. PhD thesis, TU Berlin, Inst. f. Mathematik, 2004.
- [92] G.L. Sleijpen, G.L. Booten, D.R. Fokkema, and H.A. van der Vorst. Jacobi-Davidson type methods for generalized eigenproblems and polynomial eigenproblems. *BIT*, 36:595 – 633, 1996.
- [93] G.L. Sleijpen and H.A. van der Vorst. The Jacobi–Davidson method for eigenproblems and its relation with accelerated inexact newton schemes. In S.D. Margenov and P.S. Vassilevski, editors, *Iterative Methods in Linear Algebra*, volume 3 of *IMACS Series in Computational and Applied Mathematics*, pages 377–389, New Brunswick, 1996. IMACS.
- [94] G.L. Sleijpen and H.A. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matr. Anal. Appl.*, 17:401 – 425, 1996.
- [95] G.L. Sleijpen, H.A. van der Vorst, and M. van Gijzen. Quadratic eigenproblems are no problem. *SIAM News*, 8:9–10, September 1996.
- [96] S.I. Solov’ev. Eigenvibrations of a plate with elastically attached loads. Preprint SFB393/03-06, Sonderforschungsbereich 393, Technische Universität, D-09107 Chemnitz, Germany, 2003. Available at <http://www.tu-chemnitz.de/sfb393/Files/PS/sfb03-06.ps.gz>.
- [97] D.C. Sorensen. Implicit application of polynomial filters in a  $k$ -step Arnoldi method. *SIAM J. Matrix Anal. Appl.*, 13:357–385, 1992.

- [98] M.M. Tiller. *Introduction to Physical Modeling with Modelica*. Kluwer Academic Publishers, Boston, 2nd edition, 2004.
- [99] F. Tisseur. Backward error analysis of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309:339–361, 2000.
- [100] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43:234–286, 2001.
- [101] M.B. van Gijzen. The parallel computation of the smallest eigenpair of an acoustic problem with damping. *Int. J. Numer. Meth. Engng.*, 45:765–777, 1999.
- [102] H. Voss. An error bound for eigenvalue analysis by nodal condensation. In J. Albrecht, L. Collatz, and W. Velte, editors, *Numerical Treatment of Eigenvalue Problems, Vol. 3*, volume 69 of *International Series on Numerical Mathematics*, pages 205–214, Basel, 1984. Birkhäuser.
- [103] H. Voss. A new justification of finite dynamic element methods. In J. Albrecht, L. Collatz, W. Velte, and W. Wunderlich, editors, *Numerical Treatment of Eigenvalue Problems, Vol. 4*, ISNM 83, pages 232 – 242, Basel, 1987. Birkhäuser.
- [104] H. Voss. Free vibration analysis by finite dynamic element methods. In I. Marek, editor, *Proceedings of the Second International Symposium on Numerical Analysis*, Teubner Texte zur Mathematik 107, pages 295 – 298, Leipzig, 1988. Teubner.
- [105] H. Voss. An Arnoldi method for nonlinear symmetric eigenvalue problems. In *Online Proceedings of the SIAM Conference on Applied Linear Algebra, Williamsburg*, <http://www.siam.org/meetings/laa03/>, 2003.
- [106] H. Voss. Initializing iterative projection methods for rational symmetric eigenproblems. In *Online Proceedings of the Dagstuhl Seminar: Theoretical and Computational Aspects of Matrix Algorithms, Schloss Dagstuhl 2003*, <ftp://ftp.dagstuhl.de/pub/Proceedings/03/03421/03421.VoszHeinrich.Other.pdf>, 2003.
- [107] H. Voss. A maxmin principle for nonlinear eigenvalue problems with application to a rational spectral problem in fluid–solid vibration. *Applications of Mathematics*, 48:607 – 622, 2003.

- [108] H. Voss. A rational spectral problem in fluid–solid vibration. *Elect. Trans. Num. Anal.*, 16:94 – 106, 2003.
- [109] H. Voss. An Arnoldi method for nonlinear eigenvalue problems. *BIT Numerical Mathematics*, 44:387 – 401, 2004.
- [110] H. Voss. A Jacobi–Davidson method for nonlinear eigenproblems. In M. Buback, G.D. van Albada, P.M.A. Sloom, and J.J. Dongarra, editors, *Computational Science – ICCS 2004, 4th International Conference, Kraków, Poland, June 6–9, 2004, Proceedings, Part II*, volume 3037 of *Lecture Notes in Computer Science*, pages 34–41, Berlin, Heidelberg, New York, 2004. Springer Verlag.
- [111] H. Voss. Numerical methods for sparse nonlinear eigenproblems. In Ivo Marek, editor, *Proceedings of the XV-th Summer School on Software and Algorithms of Numerical Mathematics, Hejnice, 2003*, pages 133 – 160, University of West Bohemia, Pilsen, Czech Republic, 2004.
- [112] H. Voss and B. Werner. A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems. *Math. Meth. Appl. Sci.*, 4:415–424, 1982.
- [113] H. Voss and B. Werner. Solving sparse nonlinear eigenvalue problems. Technical Report 82/4, Inst. f. Angew. Mathematik, Universität Hamburg, 1982.
- [114] D.S. Watkins. On Hamiltonian and symplectic Lanczos processes. *Linear Algebra Appl.*, 2004. To appear.
- [115] B. Werner. *Das Spektrum von Operatorenscharen mit verallgemeinerten Rayleighquotienten*. PhD thesis, Fachbereich Mathematik, Universität Hamburg, 1970.
- [116] W.H. Yang. A method for eigenvalues of sparse  $\lambda$ -matrices. *Internat. J. Numer. Meth. Engrg.*, 19:943 – 948, 1983.
- [117] R. Zurmühl and S. Falk. *Matrizen und ihre Anwendungen, Bd. I*. Springer-Verlag, Berlin, fifth edition, 1984.
- [118] R. Zurmühl and S. Falk. *Matrizen und ihre Anwendungen, Bd. II*. Springer-Verlag, Berlin, fifth edition, 1986.